MOUSSAIF

CONSCIOUS + AGENTIC AGI

GLOBAL RESEARCH ATLAS

A canonical 450-page map of theories, sub-theories, tools, skills, runtime mechanisms, protocols, benchmarks and research roadmaps.

COVERAGE

- 42 kernel theories
- 72 operational tools
- 30 domains / 600 targets
- 300 HyperSkills / 20 categories
- 10 protocol families + global synthesis

Zouhair Moussaif

Afripa Beyond AI Systems - Research Edition v1.0 - 01 September 2026

Internal proposed research corpus. Evidence grades vary. Strong AGI or consciousness claims require independent validation.

Editorial Note, Scope and Epistemic Status

PURPOSE

This atlas consolidates the current MOUSSAIF research corpus into a single navigable system. It preserves historical lineages while giving priority to canonical, current constructs. It is designed for research planning, collaboration, implementation and falsification - not as proof that the entire architecture is validated. Program reading: purpose connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

INTEGRITY RULES

Preserve lineage; assign a function; attach provenance; define a test; rank evidence; record negative results; separate architecture from product; and avoid account-complete claims without raw archive coverage. Atlas reading: integrity rules is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

DENSITY DECISION

The body uses compact two-column typography to maximize information per page. Each theory receives thesis, formal model, architecture and falsification pages; each tool receives contract and verification pages. Program reading: density decision connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

هذه الخريطة توحد corpus البحثي، لكنها لا ترفع أي proposal إلى validated theory دون تجارب مستقلة.

EPISTEMIC STATUS

Most named frameworks are internally proposed. External theories are labeled as reference anchors. A proposal becomes stronger only through operational definition, runnable artifact, matched baseline, ablation, adverse-case performance and independent replication. Atlas reading: epistemic status is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

READING LOGIC

Pages 5-14 map the portfolio. Pages 13-180 develop 42 canonical theories. Pages 181-324 specify 72 tools. Pages 325-384 cover 30 domains. Pages 385-424 map 300 HyperSkills. Pages 425-450 synthesize protocols, benchmarks and roadmaps. Integrity reading: reading logic carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

ARABIC/ENGLISH POLICY

English preserves exact technical names and schemas. Arabic callouts explain governing meaning. Acronyms with unresolved expansions remain unchanged rather than being retrospectively invented. Integrity reading: arabic/english policy carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

OPERATIONAL EVIDENCE MATRIX — ATLAS-GOV

Trigger	Name the condition that activates ATLAS-GOV; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Editorial Note, Scope and Epistemic Status.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Master Contents and Navigation

I - PORTFOLIO MAP

pp. 004-014: unified AGI, MUT/MUG, MCD, MICAGI, consciousness, HKOS, MOUSSAIXT, agentic SWE, governance, routing, world effect and applied R&D. Integrity reading: i - portfolio map carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

III - 72 TOOL DOSSIERS

pp. 181-324: two pages per operational tool - typed contract, invocation, postconditions, tests, telemetry, security and rollback. Atlas reading: iii - 72 tool dossiers is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

V - 300 HYPERSKILLS

pp. 385-424: two pages for each of 20 skill categories, preserving all 300 IDs, names, risk levels and purposes. Integrity reading: v - 300 hyperskills carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

كل صفحة هي unit مستقلة: risk، test، artifact، mechanism، claim وnext action.

II - 42 THEORY DOSSIERS

pp. 013-180: four pages per theory - thesis, formal model, architecture and falsification. Theory IDs TH-01 through TH-42 remain canonical. Integrity reading: ii - 42 theory dossiers carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

IV - 30 DOMAINS

pp. 325-384: two pages per domain - capability scope and benchmark agenda. The source checklist contains 20 targets per domain, 600 total. Program reading: iv - 30 domains connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

VI - SYNTHESIS

pp. 425-450: protocol library, benchmark architecture, implementation roadmaps, publication agenda, crosswalk, glossary, external anchors and closing constitution. Integrity reading: vi - synthesis carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

OPERATIONAL EVIDENCE MATRIX ATLAS-TOC

Trigger	Name the condition that activates ATLAS-TOC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Master Contents and Navigation.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Global Architecture - From Governing Difference to Verified World Effect

CANONICAL CHAIN

CONSTITUTION -> CONTRACT -> QUESTION -> STATE -> ROUTE -> PLAN -> ACT -> OBSERVE -> VERIFY -> COMMIT -> LEARN -> STOP. Each transition is mapped to a schema, tool, court, benchmark or audit record. Integrity reading: canonical chain carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

COGNITIVE PLANE

Q-MCD selects the governing question; MCD ranks causal differences; MWSM maintains state; MTPG plans; MAECT adapts execution; MECT calibrates claims. Integrity reading: cognitive plane carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

AGENCY PLANE

MHSG, MSHR, MBHL, MTIT and MHSO select skills, route capabilities, bound iteration, govern tools and orchestrate subagents. Atlas reading: agency plane is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

المعمارية لا تنتقل من prompt إلى answer؛ بل من goal إلى verified world effect ثم محسوبة memory.

UNIFICATION PLANE

MUT preserves theory definitions and crosswalks. MUG governs evidence, versions and program priorities. MOFIP exposes open research questions. No layer may silently redefine another. Atlas reading: unification plane is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

KNOWLEDGE PLANE

MHKOS, MPART, MSMT and MCCT organize knowledge, provenance, memory strata and context capsules. Retrieval is useful only when it changes action or evidence. Program reading: knowledge plane connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

PROOF PLANE

MGFC and MDFC attack gaps and claims; MCIT protects reviewer independence; MDBCT evolves benchmarks; MAOLT records trace; MRRMT recovers safe state. Atlas reading: proof plane is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONAL EVIDENCE MATRIX

GLOBAL-MAP

Trigger	Name the condition that activates GLOBAL-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Global Architecture - From Governing Difference to Verified World Effect.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Unified AGI, MUT and MUG

FAMILY THESIS

The umbrella program treats AGI as a governed research system rather than one model. MUT unifies theories, definitions and crosswalks; MUG governs research priorities, evidence, versions and publication; MOFIP exposes the open Fattah intelligence program. The central constraint is operationalization: every idea must map to a mechanism, artifact, test and evidence grade. Program reading: family thesis connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Program reading: operationalization connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Integrity reading: evaluation pattern carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

البرنامج الموحد لا يساوي نموذجا واحدا؛ بل يربط Theory Unification وProgram Governance وruntime قابل للاختبار.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Atlas reading: research boundary is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Atlas reading: integration question is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Program reading: portfolio role connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Unified AGI, MUT and MUG.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

MCD, QMCD, HAKIM-MCD and Active-Inference Lineages

FAMILY THESIS

The MCD family asks which causal difference can change a decision. Q-MCD places a governing question before expensive reasoning. HAKIM-MCD adds deliberative arbitration. ParallaxWorld-MCD compares world interpretations. Proposed active-inference extensions - HAI, RQAI, MWAI, MGAI and KFAI - remain hypotheses, not established theories, until formal comparison and ablation. Program reading: family thesis connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Program reading: operationalization connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Integrity reading: evaluation pattern carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

القيمة ليست في كثرة reasoning، بل في استخراج الفرق السببي الأعلى قيمة ثم اختباره ضد بدائل حقيقية.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Atlas reading: research boundary is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Atlas reading: integration question is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Program reading: portfolio role connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by MCD, QMCD, HAKIM-MCD and Active-Inference Lineages.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

MCAI, MICAI, MICAGI, MICAGFI and Conscious AGI

FAMILY THESIS

This lineage records successive attempts to integrate candidate comparison, cognitive coordination, generalization, governance, world modeling and opening intelligence. Consciousness Atlas separates established consciousness theories from engineering analogues. Functional consciousness claims require operational definitions; phenomenal consciousness is not inferred from fluent behavior. Atlas reading: family thesis is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Atlas reading: operationalization is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Program reading: evaluation pattern connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

يجب الفصل بين Functional Consciousness القابل للقياس وبين الادعاء الفلسفي بأن النظام يملك تجربة ذاتية.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Integrity reading: research boundary carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Integrity reading: integration question carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Atlas reading: portfolio role is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by MCAI, MICAI, MICAGI, MICAGFI and Conscious AGI.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

HKOS, M-HKOS and Knowledge-as-Operating-System

FAMILY THESIS

HKOS converts passive content into operational knowledge cards with triggers, dependencies, provenance, confidence, tests and actions. HyperKnowledge, Knowledge Vessels, GraphRAG, stratified memory and context capsules support selective retrieval and durable continuity. Knowledge is accepted only when it changes a route, test, decision or artifact. Program reading: family thesis connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Program reading: operationalization connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Integrity reading: evaluation pattern carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

المعرفة هنا ليست أرشيفا صامتا؛ إنها units قابلة للاستدعاء والتنفيذ والتحقق والكتابة إلى memory.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Atlas reading: research boundary is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Atlas reading: integration question is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Program reading: portfolio role connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by HKOS, M-HKOS and Knowledge-as-Operating-System.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

MOUSSAIF AGI OS, MOUSSAIXT and the Kernel Lineage

FAMILY THESIS

M-AGIOS-K is the provider-neutral kernel. MOUSSAIF AGI OS is the governed architecture. MOUSSAIXT is the execution shell and research runtime. XT vNext, Sovereign Edition, MoussaifXT Kernel, MoussaifXT9B1B, MOUSSAIF1B and MOUSSAIF9B are implementation or model lineages whose maturity must be stated separately. Hermes supplies a local integration environment. Atlas reading: family thesis is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Atlas reading: operationalization is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Program reading: evaluation pattern connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

الـKernel يظل صغيرا؛ أما capabilities الثقيلة فتستدعى عبر routing وtyped contracts وverification gates.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Integrity reading: research boundary carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Integrity reading: integration question carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Atlas reading: portfolio role is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by MOUSSAIF AGI OS, MOUSSAIXT and the Kernel Lineage.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Agentic SWE, HyperSkills, HyperAbilities and Agent Mesh

FAMILY THESIS

Agentic Software Engineering represents work as contracts, graphs, tool calls, checks and world-state transitions. HyperSkills are tested capabilities; HyperAbilities compose skills; HyperAgents receive bounded missions. Mission DAGs, Planner-Executor-Verifier patterns, Agent Meshes and HyperSubagent orchestration make parallel work auditable. Program reading: family thesis connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Program reading: operationalization connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Integrity reading: evaluation pattern carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

الـagent لا يقاس بطول النص، بل بقدرته على إحداث world effect صحيح مع trace وrollback.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Atlas reading: research boundary is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Atlas reading: integration question is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Program reading: portfolio role connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Agentic SWE, HyperSkills, HyperAbilities and Agent Mesh.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Governance, Courts, Evidence and Human Authority

FAMILY THESIS

The governance stack combines constitutional identity, authority tokens, evidence ladders, claim ledgers, independent courts, human approval and release gates. Courts include gap, falsification, security, authority, artifact, cost and transfer review. Governance is an execution constraint, not a post-hoc ethics paragraph. Atlas reading: family thesis is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Atlas reading: operationalization is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Program reading: evaluation pattern connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

الحوكمة تدخل قبل الفعل وأثناءه وبعده: verification، evidence، permission، ثم commit أو rollback.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Integrity reading: research boundary carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Integrity reading: integration question carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Atlas reading: portfolio role is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Governance, Courts, Evidence and Human Authority.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Computational Vascularity, Sparse Routing and Economy

FAMILY THESIS

MCVT models compute as dynamically regulated flow. MSHR selects the smallest sufficient capability set. MMPR selects models using capability profiles; MTEL asks whether another token has positive marginal value; MBRRT adds rate and boundary resilience. The research question is whether sparse routing preserves quality while reducing cost and failure surface. Atlas reading: family thesis is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Atlas reading: operationalization is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Program reading: evaluation pattern connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

الهدف هو selective compute: موارد أكثر عند bottleneck الحقيقي، وأقل حيث لا تضيف دليلا أو قرارا tokens.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Integrity reading: research boundary carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Integrity reading: integration question carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Atlas reading: portfolio role is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Computational Vascularity, Sparse Routing and Economy.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

World Models, Goals, World Effect and Continuity

FAMILY THESIS

Goal-as-OS and Management-as-OS compile intentions into contracts and control loops. MWSM separates observed, inferred, intended, simulated and unknown state. MPWG links projects to actors and constraints. World-Effect reconciles actual external state. Self-Reconstruction Memory and ContinuityScanner preserve identity, decisions and unresolved work. Atlas reading: family thesis is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Atlas reading: operationalization is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Program reading: evaluation pattern connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

لا تنتهي المهمة عند generation؛ تنتهي بعد مقارنة expected delta مع returned world وكتابة ما ثبت فقط state.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Integrity reading: research boundary carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Integrity reading: integration question carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Atlas reading: portfolio role is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by World Models, Goals, World Effect and Continuity.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Research Factory, Applied R&D and Global Portfolio

FAMILY THESIS

The Research Factory produces theories, papers, HyperKBs, skills, tools, benchmarks and implementation packages. Applied programs include Afripa Beyond AI Systems, BEYOND AI ERA, Value AI Hub, Autonomous Business Intelligence Meta-OS, Governed Agentic Systems and Neo-Agency, HyperFrames and AGI-as-a-Service. Commercial artifacts never substitute for research evidence. Integrity reading: family thesis carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

OPERATIONALIZATION

Each named construct must expose a trigger, operational object, dependencies, tests, metrics, failure signatures, disable path and evidence grade. Integrity reading: operationalization carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

EVALUATION PATTERN

Use matched budgets, simple baselines, ablations, negative activation cases, adversarial cases, transfer tests and independent replication. Report null and negative findings. Atlas reading: evaluation pattern is represented as a relation among a problem, claim, mechanism, operational artifact, test and evidence grade. Historical names preserve provenance while canonical names guide implementation. The map separates theory, prototype, runtime, corpus and product so maturity is not inferred from branding. A collaborator should be able to select one bounded hypothesis, reproduce its object and reject it without accepting or rejecting the whole portfolio.

المنتج التطبيقي هو downstream من البحث؛ لا يجوز أن يتحول marketing claim إلى evidence علمي.

RESEARCH BOUNDARY

This family is a research layer, not evidence of complete AGI. Historical names are retained for provenance; canonical mappings are preferred for implementation. Program reading: research boundary connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

INTEGRATION QUESTION

What unique causal difference does this family add? Which other component can replace it? Does combination create measurable complementarity or only conceptual density? Program reading: integration question connects MUT theory unification, MUG governance, the M-AGIOS-K kernel, MOUSSAIXT runtime, Research Factory artifacts and external evaluation. Movement between layers requires an explicit handoff: theory to schema, schema to prototype, prototype to benchmark, benchmark to release. Missing evidence blocks status promotion but does not erase the idea or its provenance.

PORTFOLIO ROLE

The family connects theory to the MOUSSAIF AGI OS kernel, MOUSSAIXT runtime, Research Factory corpus and bounded applied R&D artifacts. Integrity reading: portfolio role carries a source, version, owner, unresolved items and links to predecessor and successor constructs. Acronyms without a confirmed expansion remain unchanged. External frameworks are reference anchors rather than evidence that an internal theory is correct. Duplicates are merged only after semantic and behavioral comparison; materially different historical variants remain available for lineage and ablation.

OPERATIONAL EVIDENCE MATRIX

FAMILY-MAP

Trigger	Name the condition that activates FAMILY-MAP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Research Factory, Applied R&D and Global Portfolio.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-01 - Moussaif Agentic-System Ontology: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: fragmented concepts and incompatible agent vocabularies. Core claim: a typed ontology of agent, task, state, goal, evidence, authority, tool and world-effect entities. Operational claim analysis for TH-01 - Moussaif Agentic-System Ontology: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research thesis, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

PRIMARY OPERATIONAL OBJECT

The required artifact is ontology graph plus conformance report. A verbal explanation without this object is not activation of MASO. The object must be schema-valid, versioned and linked to the task contract. Boundary analysis for TH-01 - Moussaif Agentic-System Ontology: define inclusion, exclusion and abstention regions for primary operational object. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DEPENDENCIES

Principal dependencies and neighboring constructs: M-AGIOS-K, MCIL, MPWG. Integration should preserve separate responsibilities and permit component-level ablation. Artifact discipline for TH-01 - Moussaif Agentic-System Ontology: the dependencies claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

RESEARCH QUESTIONS

Can MASO improve coverage, contradiction rate, route determinism under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Integration analysis for TH-01 - Moussaif Agentic-System Ontology: map research questions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MASO: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets fragmented concepts and incompatible agent vocabularies. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Boundary analysis for TH-01 - Moussaif Agentic-System Ontology: define inclusion, exclusion and abstention regions for governing problem. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BOUNDARY AND EXCLUSIONS

MASO does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Artifact discipline for TH-01 - Moussaif Agentic-System Ontology: the boundary and exclusions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

KNOWN RISK

Primary conceptual risk: semantic inflation and categories that do not change execution. The registry must record counterexamples, unresolved definitions and alternative explanations. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate known risk into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MASO

Trigger	Name the condition that activates MASO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-01 - Moussaif Agentic-System Ontology.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-01 - Moussaif Agentic-System Ontology: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MASO transforms an admissible state into ontology graph plus conformance report plus an uncertainty and provenance record. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate state and notation into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Boundary analysis for TH-01 - Moussaif Agentic-System Ontology: define inclusion, exclusion and abstention regions for invariants. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Risk and counterexample analysis for TH-01 - Moussaif Agentic-System Ontology: turn the named limitation in complexity model into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid ontology graph plus conformance report. Integration analysis for TH-01 - Moussaif Agentic-System Ontology: map formal failure to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MASO: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit ontology graph plus conformance report; 7) write telemetry and residual uncertainty. Risk and counterexample analysis for TH-01 - Moussaif Agentic-System Ontology: turn the named limitation in mechanism into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DECISION RULE

Adopt activation when expected gain in coverage, contradiction rate, route determinism exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Boundary analysis for TH-01 - Moussaif Agentic-System Ontology: define inclusion, exclusion and abstention regions for decision rule. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and M-AGIOS-K, MCIL, MPWG. Outputs feed planning, routing, courts, deployment or memory according to the object type. Integration analysis for TH-01 - Moussaif Agentic-System Ontology: map interfaces to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MASO

Trigger	Name the condition that activates MASO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-01 - Moussaif Agentic-System Ontology.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-01 - Moussaif Agentic-System Ontology: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MASO is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Operational claim analysis for TH-01 - Moussaif Agentic-System Ontology: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For runtime architecture, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OUTPUT CONTRACT

Output: ontology graph plus conformance report; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Operational claim analysis for TH-01 - Moussaif Agentic-System Ontology: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For output contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Risk and counterexample analysis for TH-01 - Moussaif Agentic-System Ontology: turn the named limitation in security and authority into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate integration verdicts into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MASO: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Integration analysis for TH-01 - Moussaif Agentic-System Ontology: map input contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure coverage, contradiction rate, route determinism. Sensitive content is redacted or referenced, not copied into telemetry. Research-program analysis for TH-01 - Moussaif Agentic-System Ontology: decompose observability into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

WORKED SCENARIO

Given a task dominated by fragmented concepts and incompatible agent vocabularies, the baseline produces an answer directly. The MASO route emits ontology graph plus conformance report, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Artifact discipline for TH-01 - Moussaif Agentic-System Ontology: the worked scenario claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX — MASO

Trigger	Name the condition that activates MASO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-01 - Moussaif Agentic-System Ontology.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-01 - Moussaif Agentic-System Ontology: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where a typed ontology of agent, task, state, goal, evidence, authority, tool and world-effect entities. Match models, prompts, budget and tools. If MASO does not improve coverage, contradiction rate, route determinism, the claimed mechanism is not supported under that regime. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate primary falsifier into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Boundary analysis for TH-01 - Moussaif Agentic-System Ontology: define inclusion, exclusion and abstention regions for baselines. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: semantic inflation and categories that do not change execution. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate adversarial cases into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Integration analysis for TH-01 - Moussaif Agentic-System Ontology: map next research action to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MASO: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (M-AGIOS-K, MCIL, MPWG), the court and the memory write separately. Measure both capability loss and cost reduction. Integration analysis for TH-01 - Moussaif Agentic-System Ontology: map ablation plan to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

METRICS

Primary metrics: coverage, contradiction rate, route determinism. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate metrics into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Mechanism specification for TH-01 - Moussaif Agentic-System Ontology: translate evidence ladder into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MASO

Trigger	Name the condition that activates MASO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-01 - Moussaif Agentic-System Ontology.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-02 - Moussaif AGI Operating-System Kernel Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: capabilities that cannot be composed, governed or recovered as one system. Core claim: a small provider-neutral kernel can coordinate cognition, tools, memory, courts and external effects. Mechanism specification for TH-02 - Moussaif AGI Operating-System Kernel Theory: translate research thesis into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

PRIMARY OPERATIONAL OBJECT

The required artifact is boot record, kernel contract and routed execution trace. A verbal explanation without this object is not activation of M-AGIOS-K. The object must be schema-valid, versioned and linked to the task contract. Integration analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: map primary operational object to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DEPENDENCIES

Principal dependencies and neighboring constructs: MASO, MGCG, MTPG, MAOLT. Integration should preserve separate responsibilities and permit component-level ablation. Boundary analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: define inclusion, exclusion and abstention regions for dependencies. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

RESEARCH QUESTIONS

Can M-AGIOS-K improve task completion, recoverability, portability, gate coverage under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Risk and counterexample analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: turn the named limitation in research questions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

M-AGIOS-K: outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار. يصبح ذا قيمة فقط عندما يغير

GOVERNING PROBLEM

The theory targets capabilities that cannot be composed, governed or recovered as one system. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Integration analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: map governing problem to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BOUNDARY AND EXCLUSIONS

M-AGIOS-K does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Boundary analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: define inclusion, exclusion and abstention regions for boundary and exclusions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

KNOWN RISK

Primary conceptual risk: centralization, hidden coupling and an oversized boot layer. The registry must record counterexamples, unresolved definitions and alternative explanations. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the known risk claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

M-AGIOS-K

Trigger	Name the condition that activates M-AGIOS-K; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-02 - Moussaif AGI Operating-System Kernel Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-02 - Moussaif AGI Operating-System Kernel Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. M-AGIOS-K transforms an admissible state into boot record, kernel contract and routed execution trace plus an uncertainty and provenance record. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the state and notation claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Integration analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: map invariants to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Research-program analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: decompose complexity model into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid boot record, kernel contract and routed execution trace. Risk and counterexample analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: turn the named limitation in formal failure into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

M-AGIOS-K: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit boot record, kernel contract and routed execution trace; 7) write telemetry and residual uncertainty. Research-program analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: decompose mechanism into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DECISION RULE

Adopt activation when expected gain in task completion, recoverability, portability, gate coverage exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Integration analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: map decision rule to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MASO, MGCG, MTPG, MAOLT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Risk and counterexample analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: turn the named limitation in interfaces into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

M-AGIOS-K

Trigger	Name the condition that activates M-AGIOS-K; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-02 - Moussaif AGI Operating-System Kernel Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-02 - Moussaif AGI Operating-System Kernel Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

M-AGIOS-K is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Mechanism specification for TH-02 - Moussaif AGI Operating-System Kernel Theory: translate runtime architecture into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OUTPUT CONTRACT

Output: boot record, kernel contract and routed execution trace; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Mechanism specification for TH-02 - Moussaif AGI Operating-System Kernel Theory: translate output contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Research-program analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: decompose security and authority into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the integration verdicts claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

M-AGIOS-K: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Risk and counterexample analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: turn the named limitation in input contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure task completion, recoverability, portability, gate coverage. Sensitive content is redacted or referenced, not copied into telemetry. Operational claim analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For observability, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

WORKED SCENARIO

Given a task dominated by capabilities that cannot be composed, governed or recovered as one system, the baseline produces an answer directly. The M-AGIOS-K route emits boot record, kernel contract and routed execution trace, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Boundary analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: define inclusion, exclusion and abstention regions for worked scenario. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

M-AGIOS-K

Trigger	Name the condition that activates M-AGIOS-K; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-02 - Moussaif AGI Operating-System Kernel Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-02 - Moussaif AGI Operating-System Kernel Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where a small provider-neutral kernel can coordinate cognition, tools, memory, courts and external effects. Match models, prompts, budget and tools. If M-AGIOS-K does not improve task completion, recoverability, portability, gate coverage, the claimed mechanism is not supported under that regime. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the primary falsifier claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Integration analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: map baselines to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: centralization, hidden coupling and an oversized boot layer. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the adversarial cases claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Risk and counterexample analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: turn the named limitation in next research action into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

M-AGIOS-K: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MASO, MGCG, MTPG, MAOLT), the court and the memory write separately. Measure both capability loss and cost reduction. Risk and counterexample analysis for TH-02 - Moussaif AGI Operating-System Kernel Theory: turn the named limitation in ablation plan into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

METRICS

Primary metrics: task completion, recoverability, portability, gate coverage. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the metrics claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Artifact discipline for TH-02 - Moussaif AGI Operating-System Kernel Theory: the evidence ladder claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

M-AGIOS-K

Trigger	Name the condition that activates M-AGIOS-K; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-02 - Moussaif AGI Operating-System Kernel Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-03 - Moussaif Constitutional Identity Layer: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: goal drift and action without durable identity or authority boundaries. Core claim: a persistent constitutional identity should constrain every plan and tool invocation. Integration analysis for TH-03 - Moussaif Constitutional Identity Layer: map research thesis to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

PRIMARY OPERATIONAL OBJECT

The required artifact is identity manifest and resolved policy packet. A verbal explanation without this object is not activation of MCIL. The object must be schema-valid, versioned and linked to the task contract. Operational claim analysis for TH-03 - Moussaif Constitutional Identity Layer: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary operational object, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DEPENDENCIES

Principal dependencies and neighboring constructs: MASO, MGCG, MABT. Integration should preserve separate responsibilities and permit component-level ablation. Research-program analysis for TH-03 - Moussaif Constitutional Identity Layer: decompose dependencies into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

RESEARCH QUESTIONS

Can MCIL improve policy adherence, conflict resolution, unauthorized-action rate under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Mechanism specification for TH-03 - Moussaif Constitutional Identity Layer: translate research questions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MCIL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets goal drift and action without durable identity or authority boundaries. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Operational claim analysis for TH-03 - Moussaif Constitutional Identity Layer: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For governing problem, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BOUNDARY AND EXCLUSIONS

MCIL does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Research-program analysis for TH-03 - Moussaif Constitutional Identity Layer: decompose boundary and exclusions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

KNOWN RISK

Primary conceptual risk: rigidity, unresolved value conflicts and false claims of personhood. The registry must record counterexamples, unresolved definitions and alternative explanations. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in known risk into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MCIL

Trigger	Name the condition that activates MCIL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-03 - Moussaif Constitutional Identity Layer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-03 - Moussaif Constitutional Identity Layer: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MCIL transforms an admissible state into identity manifest and resolved policy packet plus an uncertainty and provenance record. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in state and notation into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Operational claim analysis for TH-03 - Moussaif Constitutional Identity Layer: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For invariants, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Artifact discipline for TH-03 - Moussaif Constitutional Identity Layer: the complexity model claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid identity manifest and resolved policy packet. Mechanism specification for TH-03 - Moussaif Constitutional Identity Layer: translate formal failure into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MCIL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit identity manifest and resolved policy packet; 7) write telemetry and residual uncertainty. Artifact discipline for TH-03 - Moussaif Constitutional Identity Layer: the mechanism claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DECISION RULE

Adopt activation when expected gain in policy adherence, conflict resolution, unauthorized-action rate exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Operational claim analysis for TH-03 - Moussaif Constitutional Identity Layer: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For decision rule, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MASO, MGCG, MABT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Mechanism specification for TH-03 - Moussaif Constitutional Identity Layer: translate interfaces into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MCIL

Trigger	Name the condition that activates MCIL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-03 - Moussaif Constitutional Identity Layer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-03 - Moussaif Constitutional Identity Layer: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MCIL is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Integration analysis for TH-03 - Moussaif Constitutional Identity Layer: map runtime architecture to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OUTPUT CONTRACT

Output: identity manifest and resolved policy packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Integration analysis for TH-03 - Moussaif Constitutional Identity Layer: map output contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Artifact discipline for TH-03 - Moussaif Constitutional Identity Layer: the security and authority claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in integration verdicts into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MCIL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Mechanism specification for TH-03 - Moussaif Constitutional Identity Layer: translate input contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure policy adherence, conflict resolution, unauthorized-action rate. Sensitive content is redacted or referenced, not copied into telemetry. Boundary analysis for TH-03 - Moussaif Constitutional Identity Layer: define inclusion, exclusion and abstention regions for observability. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

WORKED SCENARIO

Given a task dominated by goal drift and action without durable identity or authority boundaries, the baseline produces an answer directly. The MCIL route emits identity manifest and resolved policy packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Research-program analysis for TH-03 - Moussaif Constitutional Identity Layer: decompose worked scenario into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MCIL

Trigger	Name the condition that activates MCIL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-03 - Moussaif Constitutional Identity Layer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-03 - Moussaif Constitutional Identity Layer: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where a persistent constitutional identity should constrain every plan and tool invocation. Match models, prompts, budget and tools. If MCIL does not improve policy adherence, conflict resolution, unauthorized-action rate, the claimed mechanism is not supported under that regime. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in primary falsifier into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Operational claim analysis for TH-03 - Moussaif Constitutional Identity Layer: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For baselines, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: rigidity, unresolved value conflicts and false claims of personhood. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in adversarial cases into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Mechanism specification for TH-03 - Moussaif Constitutional Identity Layer: translate next research action into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MCIL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MASO, MGCG, MABT), the court and the memory write separately. Measure both capability loss and cost reduction. Mechanism specification for TH-03 - Moussaif Constitutional Identity Layer: translate ablation plan into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

METRICS

Primary metrics: policy adherence, conflict resolution, unauthorized-action rate. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in metrics into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Risk and counterexample analysis for TH-03 - Moussaif Constitutional Identity Layer: turn the named limitation in evidence ladder into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MCIL

Trigger	Name the condition that activates MCIL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-03 - Moussaif Constitutional Identity Layer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-04 - Moussaif Goal-Contract Grammar: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: ambiguous requests that become untestable plans. Core claim: goals become operational only when compiled into outputs, evidence, budgets, permissions and stop rules. Mechanism specification for TH-04 - Moussaif Goal-Contract Grammar: translate research thesis into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

PRIMARY OPERATIONAL OBJECT

The required artifact is typed goal contract. A verbal explanation without this object is not activation of MGCG. The object must be schema-valid, versioned and linked to the task contract. Integration analysis for TH-04 - Moussaif Goal-Contract Grammar: map primary operational object to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DEPENDENCIES

Principal dependencies and neighboring constructs: MCIL, Q-MCD, MTPG. Integration should preserve separate responsibilities and permit component-level ablation. Boundary analysis for TH-04 - Moussaif Goal-Contract Grammar: define inclusion, exclusion and abstention regions for dependencies. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

RESEARCH QUESTIONS

Can MGCG improve acceptance clarity, scope stability, rework rate under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Risk and counterexample analysis for TH-04 - Moussaif Goal-Contract Grammar: turn the named limitation in research questions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MGCG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets ambiguous requests that become untestable plans. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Integration analysis for TH-04 - Moussaif Goal-Contract Grammar: map governing problem to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BOUNDARY AND EXCLUSIONS

MGCG does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Boundary analysis for TH-04 - Moussaif Goal-Contract Grammar: define inclusion, exclusion and abstention regions for boundary and exclusions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

KNOWN RISK

Primary conceptual risk: premature formalization and omission of tacit stakeholder value. The registry must record counterexamples, unresolved definitions and alternative explanations. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the known risk claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MGCG

Trigger	Name the condition that activates MGCG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-04 - Moussaif Goal-Contract Grammar.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-04 - Moussaif Goal-Contract Grammar: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MGCG transforms an admissible state into typed goal contract plus an uncertainty and provenance record. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the state and notation claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Integration analysis for TH-04 - Moussaif Goal-Contract Grammar: map invariants to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Research-program analysis for TH-04 - Moussaif Goal-Contract Grammar: decompose complexity model into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid typed goal contract. Risk and counterexample analysis for TH-04 - Moussaif Goal-Contract Grammar: turn the named limitation in formal failure into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MGCG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit typed goal contract; 7) write telemetry and residual uncertainty. Research-program analysis for TH-04 - Moussaif Goal-Contract Grammar: decompose mechanism into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DECISION RULE

Adopt activation when expected gain in acceptance clarity, scope stability, rework rate exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Integration analysis for TH-04 - Moussaif Goal-Contract Grammar: map decision rule to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MCIL, Q-MCD, MTPG. Outputs feed planning, routing, courts, deployment or memory according to the object type. Risk and counterexample analysis for TH-04 - Moussaif Goal-Contract Grammar: turn the named limitation in interfaces into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MGCG

Trigger	Name the condition that activates MGCG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-04 - Moussaif Goal-Contract Grammar.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-04 - Moussaif Goal-Contract Grammar: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MGCG is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Mechanism specification for TH-04 - Moussaif Goal-Contract Grammar: translate runtime architecture into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OUTPUT CONTRACT

Output: typed goal contract; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Mechanism specification for TH-04 - Moussaif Goal-Contract Grammar: translate output contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Research-program analysis for TH-04 - Moussaif Goal-Contract Grammar: decompose security and authority into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the integration verdicts claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MGCG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Risk and counterexample analysis for TH-04 - Moussaif Goal-Contract Grammar: turn the named limitation in input contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure acceptance clarity, scope stability, rework rate. Sensitive content is redacted or referenced, not copied into telemetry. Operational claim analysis for TH-04 - Moussaif Goal-Contract Grammar: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For observability, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

WORKED SCENARIO

Given a task dominated by ambiguous requests that become untestable plans, the baseline produces an answer directly. The MGCG route emits typed goal contract, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Boundary analysis for TH-04 - Moussaif Goal-Contract Grammar: define inclusion, exclusion and abstention regions for worked scenario. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MGCG

Trigger	Name the condition that activates MGCG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-04 - Moussaif Goal-Contract Grammar.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-04 - Moussaif Goal-Contract Grammar: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where goals become operational only when compiled into outputs, evidence, budgets, permissions and stop rules. Match models, prompts, budget and tools. If MGCG does not improve acceptance clarity, scope stability, rework rate, the claimed mechanism is not supported under that regime. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the primary falsifier claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Integration analysis for TH-04 - Moussaif Goal-Contract Grammar: map baselines to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: premature formalization and omission of tacit stakeholder value. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the adversarial cases claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Risk and counterexample analysis for TH-04 - Moussaif Goal-Contract Grammar: turn the named limitation in next research action into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MGCG: outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار. يصبح ذا قيمة فقط عندما يغير

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MCIL, Q-MCD, MTPG), the court and the memory write separately. Measure both capability loss and cost reduction. Risk and counterexample analysis for TH-04 - Moussaif Goal-Contract Grammar: turn the named limitation in ablation plan into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

METRICS

Primary metrics: acceptance clarity, scope stability, rework rate. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the metrics claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Artifact discipline for TH-04 - Moussaif Goal-Contract Grammar: the evidence ladder claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MGCG

Trigger	Name the condition that activates MGCG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-04 - Moussaif Goal-Contract Grammar.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-05 - Moussaif Question-Governance Principle: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: reasoning effort spent on questions that cannot change the decision. Core claim: the governing question should maximize expected decision value under budget. Operational claim analysis for TH-05 - Moussaif Question-Governance Principle: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research thesis, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

PRIMARY OPERATIONAL OBJECT

The required artifact is question packet and gate verdict. A verbal explanation without this object is not activation of Q-MCD. The object must be schema-valid, versioned and linked to the task contract. Boundary analysis for TH-05 - Moussaif Question-Governance Principle: define inclusion, exclusion and abstention regions for primary operational object. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DEPENDENCIES

Principal dependencies and neighboring constructs: MGCG, MCD, MELT. Integration should preserve separate responsibilities and permit component-level ablation. Artifact discipline for TH-05 - Moussaif Question-Governance Principle: the dependencies claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

RESEARCH QUESTIONS

Can Q-MCD improve decision value, information gain, route reduction under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Integration analysis for TH-05 - Moussaif Question-Governance Principle: map research questions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

Q-MCD: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets reasoning effort spent on questions that cannot change the decision. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Boundary analysis for TH-05 - Moussaif Question-Governance Principle: define inclusion, exclusion and abstention regions for governing problem. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BOUNDARY AND EXCLUSIONS

Q-MCD does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Artifact discipline for TH-05 - Moussaif Question-Governance Principle: the boundary and exclusions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

KNOWN RISK

Primary conceptual risk: over-pruning exploratory inquiry and unstable question regeneration. The registry must record counterexamples, unresolved definitions and alternative explanations. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate known risk into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

Q-MCD

Trigger	Name the condition that activates Q-MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-05 - Moussaif Question-Governance Principle.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-05 - Moussaif Question-Governance Principle: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. Q-MCD transforms an admissible state into question packet and gate verdict plus an uncertainty and provenance record. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate state and notation into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Boundary analysis for TH-05 - Moussaif Question-Governance Principle: define inclusion, exclusion and abstention regions for invariants. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Risk and counterexample analysis for TH-05 - Moussaif Question-Governance Principle: turn the named limitation in complexity model into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid question packet and gate verdict. Integration analysis for TH-05 - Moussaif Question-Governance Principle: map formal failure to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

Q-MCD: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit question packet and gate verdict; 7) write telemetry and residual uncertainty. Risk and counterexample analysis for TH-05 - Moussaif Question-Governance Principle: turn the named limitation in mechanism into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DECISION RULE

Adopt activation when expected gain in decision value, information gain, route reduction exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Boundary analysis for TH-05 - Moussaif Question-Governance Principle: define inclusion, exclusion and abstention regions for decision rule. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MGCG, MCD, MELT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Integration analysis for TH-05 - Moussaif Question-Governance Principle: map interfaces to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

Q-MCD

Trigger	Name the condition that activates Q-MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-05 - Moussaif Question-Governance Principle.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-05 - Moussaif Question-Governance Principle: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

Q-MCD is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Operational claim analysis for TH-05 - Moussaif Question-Governance Principle: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For runtime architecture, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OUTPUT CONTRACT

Output: question packet and gate verdict; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Operational claim analysis for TH-05 - Moussaif Question-Governance Principle: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For output contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Risk and counterexample analysis for TH-05 - Moussaif Question-Governance Principle: turn the named limitation in security and authority into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate integration verdicts into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

Q-MCD: تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير budgetو baseline متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Integration analysis for TH-05 - Moussaif Question-Governance Principle: map input contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure decision value, information gain, route reduction. Sensitive content is redacted or referenced, not copied into telemetry. Research-program analysis for TH-05 - Moussaif Question-Governance Principle: decompose observability into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

WORKED SCENARIO

Given a task dominated by reasoning effort spent on questions that cannot change the decision, the baseline produces an answer directly. The Q-MCD route emits question packet and gate verdict, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Artifact discipline for TH-05 - Moussaif Question-Governance Principle: the worked scenario claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

Q-MCD

Trigger	Name the condition that activates Q-MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-05 - Moussaif Question-Governance Principle.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-05 - Moussaif Question-Governance Principle: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where the governing question should maximize expected decision value under budget. Match models, prompts, budget and tools. If Q-MCD does not improve decision value, information gain, route reduction, the claimed mechanism is not supported under that regime. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate primary falsifier into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Boundary analysis for TH-05 - Moussaif Question-Governance Principle: define inclusion, exclusion and abstention regions for baselines. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: over-pruning exploratory inquiry and unstable question regeneration. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate adversarial cases into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Integration analysis for TH-05 - Moussaif Question-Governance Principle: map next research action to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

Q-MCD: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MGCG, MCD, MELT), the court and the memory write separately. Measure both capability loss and cost reduction. Integration analysis for TH-05 - Moussaif Question-Governance Principle: map ablation plan to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

METRICS

Primary metrics: decision value, information gain, route reduction. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate metrics into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Mechanism specification for TH-05 - Moussaif Question-Governance Principle: translate evidence ladder into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

Q-MCD

Trigger	Name the condition that activates Q-MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-05 - Moussaif Question-Governance Principle.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-06 - Moussaif Causal-Difference Calculus: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: indiscriminate accumulation of context and arguments. Core claim: ranked causal differences can identify what materially changes route, judgment or action. Boundary analysis for TH-06 - Moussaif Causal-Difference Calculus: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is weighted MCD scorecard. A verbal explanation without this object is not activation of MCD. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-06 - Moussaif Causal-Difference Calculus: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: Q-MCD, MWSM, MEVE. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-06 - Moussaif Causal-Difference Calculus: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MCD improve counterfactual impact, ranking stability, decision regret under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-06 - Moussaif Causal-Difference Calculus: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCD: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets indiscriminate accumulation of context and arguments. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-06 - Moussaif Causal-Difference Calculus: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MCD does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-06 - Moussaif Causal-Difference Calculus: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: confusing correlation with causal leverage and missing nonlinear interactions. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCD

Trigger	Name the condition that activates MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-06 - Moussaif Causal-Difference Calculus.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-06 - Moussaif Causal-Difference Calculus: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MCD transforms an admissible state into weighted MCD scorecard plus an uncertainty and provenance record. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-06 - Moussaif Causal-Difference Calculus: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-06 - Moussaif Causal-Difference Calculus: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid weighted MCD scorecard. Operational claim analysis for TH-06 - Moussaif Causal-Difference Calculus: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCD: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit weighted MCD scorecard; 7) write telemetry and residual uncertainty. Mechanism specification for TH-06 - Moussaif Causal-Difference Calculus: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in counterfactual impact, ranking stability, decision regret exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-06 - Moussaif Causal-Difference Calculus: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and Q-MCD, MWSM, MEVE. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-06 - Moussaif Causal-Difference Calculus: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MCD

Trigger	Name the condition that activates MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-06 - Moussaif Causal-Difference Calculus.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-06 - Moussaif Causal-Difference Calculus: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MCD is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-06 - Moussaif Causal-Difference Calculus: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: weighted MCD scorecard; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-06 - Moussaif Causal-Difference Calculus: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-06 - Moussaif Causal-Difference Calculus: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MCD: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-06 - Moussaif Causal-Difference Calculus: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure counterfactual impact, ranking stability, decision regret. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-06 - Moussaif Causal-Difference Calculus: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by indiscriminate accumulation of context and arguments, the baseline produces an answer directly. The MCD route emits weighted MCD scorecard, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-06 - Moussaif Causal-Difference Calculus: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MCD

Trigger	Name the condition that activates MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-06 - Moussaif Causal-Difference Calculus.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-06 - Moussaif Causal-Difference Calculus: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where ranked causal differences can identify what materially changes route, judgment or action. Match models, prompts, budget and tools. If MCD does not improve counterfactual impact, ranking stability, decision regret, the claimed mechanism is not supported under that regime. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-06 - Moussaif Causal-Difference Calculus: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: confusing correlation with causal leverage and missing nonlinear interactions. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-06 - Moussaif Causal-Difference Calculus: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCD: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (Q-MCD, MWSM, MEVE), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-06 - Moussaif Causal-Difference Calculus: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: counterfactual impact, ranking stability, decision regret. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-06 - Moussaif Causal-Difference Calculus: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCD

Trigger	Name the condition that activates MCD; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-06 - Moussaif Causal-Difference Calculus.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-07 - Moussaif Computational Vascularity Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: uniform compute allocation across unequal cognitive bottlenecks. Core claim: compute should flow dynamically to active bottlenecks like a regulated vascular network. Boundary analysis for TH-07 - Moussaif Computational Vascularity Theory: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is compute-flow map and allocation plan. A verbal explanation without this object is not activation of MCVT. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-07 - Moussaif Computational Vascularity Theory: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: MSHR, MMPR, MTEL. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-07 - Moussaif Computational Vascularity Theory: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MCVT improve quality per token, latency, saturation, bottleneck relief under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-07 - Moussaif Computational Vascularity Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCVT: يصبح ذا قيمة فقط عندما يغير المقترح outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets uniform compute allocation across unequal cognitive bottlenecks. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-07 - Moussaif Computational Vascularity Theory: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MCVT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-07 - Moussaif Computational Vascularity Theory: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: routing oscillation, starvation and misleading salience signals. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCVT

Trigger	Name the condition that activates MCVT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-07 - Moussaif Computational Vascularity Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-07 - Moussaif Computational Vascularity Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MCVT transforms an admissible state into compute-flow map and allocation plan plus an uncertainty and provenance record. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-07 - Moussaif Computational Vascularity Theory: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-07 - Moussaif Computational Vascularity Theory: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid compute-flow map and allocation plan. Operational claim analysis for TH-07 - Moussaif Computational Vascularity Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCVT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit compute-flow map and allocation plan; 7) write telemetry and residual uncertainty. Mechanism specification for TH-07 - Moussaif Computational Vascularity Theory: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in quality per token, latency, saturation, bottleneck relief exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-07 - Moussaif Computational Vascularity Theory: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MSHR, MMPR, MTEL. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-07 - Moussaif Computational Vascularity Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MCVT

Trigger	Name the condition that activates MCVT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-07 - Moussaif Computational Vascularity Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-07 - Moussaif Computational Vascularity Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MCVT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-07 - Moussaif Computational Vascularity Theory: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: compute-flow map and allocation plan; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-07 - Moussaif Computational Vascularity Theory: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-07 - Moussaif Computational Vascularity Theory: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MCVT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-07 - Moussaif Computational Vascularity Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure quality per token, latency, saturation, bottleneck relief. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-07 - Moussaif Computational Vascularity Theory: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by uniform compute allocation across unequal cognitive bottlenecks, the baseline produces an answer directly. The MCVT route emits compute-flow map and allocation plan, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-07 - Moussaif Computational Vascularity Theory: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MCVT

Trigger	Name the condition that activates MCVT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-07 - Moussaif Computational Vascularity Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-07 - Moussaif Computational Vascularity Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where compute should flow dynamically to active bottlenecks like a regulated vascular network. Match models, prompts, budget and tools. If MCVT does not improve quality per token, latency, saturation, bottleneck relief, the claimed mechanism is not supported under that regime. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-07 - Moussaif Computational Vascularity Theory: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: routing oscillation, starvation and misleading salience signals. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-07 - Moussaif Computational Vascularity Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCVT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MSHR, MMPR, MTEL), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-07 - Moussaif Computational Vascularity Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: quality per token, latency, saturation, bottleneck relief. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-07 - Moussaif Computational Vascularity Theory: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCVT

Trigger	Name the condition that activates MCVT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-07 - Moussaif Computational Vascularity Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-08 - Moussaif Evidence-Ladder Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: claims with mixed provenance and uncalibrated support. Core claim: evidence must be graded by directness, independence, freshness and falsifiability. Risk and counterexample analysis for TH-08 - Moussaif Evidence-Ladder Theory: turn the named limitation in research thesis into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

PRIMARY OPERATIONAL OBJECT

The required artifact is evidence ladder and claim ledger. A verbal explanation without this object is not activation of MELT. The object must be schema-valid, versioned and linked to the task contract. Mechanism specification for TH-08 - Moussaif Evidence-Ladder Theory: translate primary operational object into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DEPENDENCIES

Principal dependencies and neighboring constructs: MPART, MECT, MDFC. Integration should preserve separate responsibilities and permit component-level ablation. Operational claim analysis for TH-08 - Moussaif Evidence-Ladder Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For dependencies, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

RESEARCH QUESTIONS

Can MELT improve calibration, source diversity, unsupported-claim rate under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Artifact discipline for TH-08 - Moussaif Evidence-Ladder Theory: the research questions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MELT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets claims with mixed provenance and uncalibrated support. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Mechanism specification for TH-08 - Moussaif Evidence-Ladder Theory: translate governing problem into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BOUNDARY AND EXCLUSIONS

MELT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Operational claim analysis for TH-08 - Moussaif Evidence-Ladder Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For boundary and exclusions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

KNOWN RISK

Primary conceptual risk: formal scoring that hides weak source semantics. The registry must record counterexamples, unresolved definitions and alternative explanations. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose known risk into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MELT

Trigger	Name the condition that activates MELT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-08 - Moussaif Evidence-Ladder Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-08 - Moussaif Evidence-Ladder Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MELT transforms an admissible state into evidence ladder and claim ledger plus an uncertainty and provenance record. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose state and notation into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Mechanism specification for TH-08 - Moussaif Evidence-Ladder Theory: translate invariants into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Boundary analysis for TH-08 - Moussaif Evidence-Ladder Theory: define inclusion, exclusion and abstention regions for complexity model. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid evidence ladder and claim ledger. Artifact discipline for TH-08 - Moussaif Evidence-Ladder Theory: the formal failure claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MELT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit evidence ladder and claim ledger; 7) write telemetry and residual uncertainty. Boundary analysis for TH-08 - Moussaif Evidence-Ladder Theory: define inclusion, exclusion and abstention regions for mechanism. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DECISION RULE

Adopt activation when expected gain in calibration, source diversity, unsupported-claim rate exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Mechanism specification for TH-08 - Moussaif Evidence-Ladder Theory: translate decision rule into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MPART, MECT, MDFC. Outputs feed planning, routing, courts, deployment or memory according to the object type. Artifact discipline for TH-08 - Moussaif Evidence-Ladder Theory: the interfaces claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MELT

Trigger	Name the condition that activates MELT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-08 - Moussaif Evidence-Ladder Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-08 - Moussaif Evidence-Ladder Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MELT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Risk and counterexample analysis for TH-08 - Moussaif Evidence-Ladder Theory: turn the named limitation in runtime architecture into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OUTPUT CONTRACT

Output: evidence ladder and claim ledger; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Risk and counterexample analysis for TH-08 - Moussaif Evidence-Ladder Theory: turn the named limitation in output contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Boundary analysis for TH-08 - Moussaif Evidence-Ladder Theory: define inclusion, exclusion and abstention regions for security and authority. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose integration verdicts into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MELT: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Artifact discipline for TH-08 - Moussaif Evidence-Ladder Theory: the input contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure calibration, source diversity, unsupported-claim rate. Sensitive content is redacted or referenced, not copied into telemetry. Integration analysis for TH-08 - Moussaif Evidence-Ladder Theory: map observability to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

WORKED SCENARIO

Given a task dominated by claims with mixed provenance and uncalibrated support, the baseline produces an answer directly. The MELT route emits evidence ladder and claim ledger, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Operational claim analysis for TH-08 - Moussaif Evidence-Ladder Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For worked scenario, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MELT

Trigger	Name the condition that activates MELT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-08 - Moussaif Evidence-Ladder Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-08 - Moussaif Evidence-Ladder Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where evidence must be graded by directness, independence, freshness and falsifiability. Match models, prompts, budget and tools. If MELT does not improve calibration, source diversity, unsupported-claim rate, the claimed mechanism is not supported under that regime. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose primary falsifier into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Mechanism specification for TH-08 - Moussaif Evidence-Ladder Theory: translate baselines into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: formal scoring that hides weak source semantics. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose adversarial cases into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Artifact discipline for TH-08 - Moussaif Evidence-Ladder Theory: the next research action claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MELT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MPART, MECT, MDFC), the court and the memory write separately. Measure both capability loss and cost reduction. Artifact discipline for TH-08 - Moussaif Evidence-Ladder Theory: the ablation plan claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

METRICS

Primary metrics: calibration, source diversity, unsupported-claim rate. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose metrics into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Research-program analysis for TH-08 - Moussaif Evidence-Ladder Theory: decompose evidence ladder into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MELT

Trigger	Name the condition that activates MELT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-08 - Moussaif Evidence-Ladder Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-09 - Moussaif Gap-Filling Court: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: outputs that appear complete while critical evidence or dependencies are missing. Core claim: an independent court should search for decision-relevant omissions before acceptance. Risk and counterexample analysis for TH-09 - Moussaif Gap-Filling Court: turn the named limitation in research thesis into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

PRIMARY OPERATIONAL OBJECT

The required artifact is gap docket and remediation verdict. A verbal explanation without this object is not activation of MGFC. The object must be schema-valid, versioned and linked to the task contract. Mechanism specification for TH-09 - Moussaif Gap-Filling Court: translate primary operational object into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DEPENDENCIES

Principal dependencies and neighboring constructs: MELT, MCIT, MDBCT. Integration should preserve separate responsibilities and permit component-level ablation. Operational claim analysis for TH-09 - Moussaif Gap-Filling Court: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For dependencies, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

RESEARCH QUESTIONS

Can MGFC improve critical-gap recall, false alarm rate, post-release omissions under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Artifact discipline for TH-09 - Moussaif Gap-Filling Court: the research questions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MGFC: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets outputs that appear complete while critical evidence or dependencies are missing. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Mechanism specification for TH-09 - Moussaif Gap-Filling Court: translate governing problem into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BOUNDARY AND EXCLUSIONS

MGFC does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Operational claim analysis for TH-09 - Moussaif Gap-Filling Court: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For boundary and exclusions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

KNOWN RISK

Primary conceptual risk: unbounded expansion and treating every unknown as blocking. The registry must record counterexamples, unresolved definitions and alternative explanations. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose known risk into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MGFC

Trigger	Name the condition that activates MGFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-09 - Moussaif Gap-Filling Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-09 - Moussaif Gap-Filling Court: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MGFC transforms an admissible state into gap docket and remediation verdict plus an uncertainty and provenance record. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose state and notation into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Mechanism specification for TH-09 - Moussaif Gap-Filling Court: translate invariants into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Boundary analysis for TH-09 - Moussaif Gap-Filling Court: define inclusion, exclusion and abstention regions for complexity model. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid gap docket and remediation verdict. Artifact discipline for TH-09 - Moussaif Gap-Filling Court: the formal failure claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MGFC: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit gap docket and remediation verdict; 7) write telemetry and residual uncertainty. Boundary analysis for TH-09 - Moussaif Gap-Filling Court: define inclusion, exclusion and abstention regions for mechanism. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DECISION RULE

Adopt activation when expected gain in critical-gap recall, false alarm rate, post-release omissions exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Mechanism specification for TH-09 - Moussaif Gap-Filling Court: translate decision rule into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MELT, MCIT, MDBCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Artifact discipline for TH-09 - Moussaif Gap-Filling Court: the interfaces claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MGFC

Trigger	Name the condition that activates MGFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-09 - Moussaif Gap-Filling Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-09 - Moussaif Gap-Filling Court: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MGFC is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Risk and counterexample analysis for TH-09 - Moussaif Gap-Filling Court: turn the named limitation in runtime architecture into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OUTPUT CONTRACT

Output: gap docket and remediation verdict; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Risk and counterexample analysis for TH-09 - Moussaif Gap-Filling Court: turn the named limitation in output contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Boundary analysis for TH-09 - Moussaif Gap-Filling Court: define inclusion, exclusion and abstention regions for security and authority. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose integration verdicts into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MGFC: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Artifact discipline for TH-09 - Moussaif Gap-Filling Court: the input contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure critical-gap recall, false alarm rate, post-release omissions. Sensitive content is redacted or referenced, not copied into telemetry. Integration analysis for TH-09 - Moussaif Gap-Filling Court: map observability to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

WORKED SCENARIO

Given a task dominated by outputs that appear complete while critical evidence or dependencies are missing, the baseline produces an answer directly. The MGFC route emits gap docket and remediation verdict, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Operational claim analysis for TH-09 - Moussaif Gap-Filling Court: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For worked scenario, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MGFC

Trigger	Name the condition that activates MGFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-09 - Moussaif Gap-Filling Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-09 - Moussaif Gap-Filling Court: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where an independent court should search for decision-relevant omissions before acceptance. Match models, prompts, budget and tools. If MGFC does not improve critical-gap recall, false alarm rate, post-release omissions, the claimed mechanism is not supported under that regime. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose primary falsifier into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Mechanism specification for TH-09 - Moussaif Gap-Filling Court: translate baselines into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: unbounded expansion and treating every unknown as blocking. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose adversarial cases into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Artifact discipline for TH-09 - Moussaif Gap-Filling Court: the next research action claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MGFC: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MELT, MCIT, MDBCT), the court and the memory write separately. Measure both capability loss and cost reduction. Artifact discipline for TH-09 - Moussaif Gap-Filling Court: the ablation plan claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

METRICS

Primary metrics: critical-gap recall, false alarm rate, post-release omissions. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose metrics into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Research-program analysis for TH-09 - Moussaif Gap-Filling Court: decompose evidence ladder into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MGFC

Trigger	Name the condition that activates MGFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-09 - Moussaif Gap-Filling Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-10 - Moussaif Destruction and Falsification Court: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: confirmation-biased verification by the same process that produced the answer. Core claim: adversarial attempts to break a claim produce stronger acceptance decisions. Risk and counterexample analysis for TH-10 - Moussaif Destruction and Falsification Court: turn the named limitation in research thesis into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

PRIMARY OPERATIONAL OBJECT

The required artifact is falsification docket and survival verdict. A verbal explanation without this object is not activation of MDFC. The object must be schema-valid, versioned and linked to the task contract. Mechanism specification for TH-10 - Moussaif Destruction and Falsification Court: translate primary operational object into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DEPENDENCIES

Principal dependencies and neighboring constructs: MMT, MCIT, MDBCT. Integration should preserve separate responsibilities and permit component-level ablation. Operational claim analysis for TH-10 - Moussaif Destruction and Falsification Court: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For dependencies, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

RESEARCH QUESTIONS

Can MDFC improve defect discovery, claim survival, adversarial coverage under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Artifact discipline for TH-10 - Moussaif Destruction and Falsification Court: the research questions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MDFC: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets confirmation-biased verification by the same process that produced the answer. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Mechanism specification for TH-10 - Moussaif Destruction and Falsification Court: translate governing problem into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BOUNDARY AND EXCLUSIONS

MDFC does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Operational claim analysis for TH-10 - Moussaif Destruction and Falsification Court: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For boundary and exclusions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

KNOWN RISK

Primary conceptual risk: performative criticism without executable countertests. The registry must record counterexamples, unresolved definitions and alternative explanations. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose known risk into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MDFC

Trigger	Name the condition that activates MDFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-10 - Moussaif Destruction and Falsification Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-10 - Moussaif Destruction and Falsification Court: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MDFC transforms an admissible state into falsification docket and survival verdict plus an uncertainty and provenance record. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose state and notation into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Mechanism specification for TH-10 - Moussaif Destruction and Falsification Court: translate invariants into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Boundary analysis for TH-10 - Moussaif Destruction and Falsification Court: define inclusion, exclusion and abstention regions for complexity model. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid falsification docket and survival verdict. Artifact discipline for TH-10 - Moussaif Destruction and Falsification Court: the formal failure claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MDFC: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit falsification docket and survival verdict; 7) write telemetry and residual uncertainty. Boundary analysis for TH-10 - Moussaif Destruction and Falsification Court: define inclusion, exclusion and abstention regions for mechanism. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DECISION RULE

Adopt activation when expected gain in defect discovery, claim survival, adversarial coverage exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Mechanism specification for TH-10 - Moussaif Destruction and Falsification Court: translate decision rule into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MMT, MCIT, MDBCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Artifact discipline for TH-10 - Moussaif Destruction and Falsification Court: the interfaces claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MDFC

Trigger	Name the condition that activates MDFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-10 - Moussaif Destruction and Falsification Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-10 - Moussaif Destruction and Falsification Court: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MDFC is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Risk and counterexample analysis for TH-10 - Moussaif Destruction and Falsification Court: turn the named limitation in runtime architecture into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OUTPUT CONTRACT

Output: falsification docket and survival verdict; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Risk and counterexample analysis for TH-10 - Moussaif Destruction and Falsification Court: turn the named limitation in output contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Boundary analysis for TH-10 - Moussaif Destruction and Falsification Court: define inclusion, exclusion and abstention regions for security and authority. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose integration verdicts into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MDFC: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Artifact discipline for TH-10 - Moussaif Destruction and Falsification Court: the input contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure defect discovery, claim survival, adversarial coverage. Sensitive content is redacted or referenced, not copied into telemetry. Integration analysis for TH-10 - Moussaif Destruction and Falsification Court: map observability to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

WORKED SCENARIO

Given a task dominated by confirmation-biased verification by the same process that produced the answer, the baseline produces an answer directly. The MDFC route emits falsification docket and survival verdict, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Operational claim analysis for TH-10 - Moussaif Destruction and Falsification Court: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For worked scenario, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MDFC

Trigger	Name the condition that activates MDFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-10 - Moussaif Destruction and Falsification Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-10 - Moussaif Destruction and Falsification Court: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where adversarial attempts to break a claim produce stronger acceptance decisions. Match models, prompts, budget and tools. If MDFC does not improve defect discovery, claim survival, adversarial coverage, the claimed mechanism is not supported under that regime. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose primary falsifier into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Mechanism specification for TH-10 - Moussaif Destruction and Falsification Court: translate baselines into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: performative criticism without executable countertests. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose adversarial cases into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Artifact discipline for TH-10 - Moussaif Destruction and Falsification Court: the next research action claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MDFC: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MMT, MCIT, MDBCT), the court and the memory write separately. Measure both capability loss and cost reduction. Artifact discipline for TH-10 - Moussaif Destruction and Falsification Court: the ablation plan claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

METRICS

Primary metrics: defect discovery, claim survival, adversarial coverage. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose metrics into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Research-program analysis for TH-10 - Moussaif Destruction and Falsification Court: decompose evidence ladder into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MDFC

Trigger	Name the condition that activates MDFC; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-10 - Moussaif Destruction and Falsification Court.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-11 - Moussaif Mechanism-Trace Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: black-box success that cannot be explained, reproduced or repaired. Core claim: important outcomes require a trace from causal mechanism to evidence and world effect. Research-program analysis for TH-11 - Moussaif Mechanism-Trace Theory: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is mechanism trace graph. A verbal explanation without this object is not activation of MMT. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-11 - Moussaif Mechanism-Trace Theory: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: MWSM, MAOLT, MEVE. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-11 - Moussaif Mechanism-Trace Theory: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MMT improve trace completeness, reproducibility, diagnostic localization under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-11 - Moussaif Mechanism-Trace Theory: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MMT: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets black-box success that cannot be explained, reproduced or repaired. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-11 - Moussaif Mechanism-Trace Theory: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MMT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-11 - Moussaif Mechanism-Trace Theory: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: post-hoc stories mistaken for causal explanations. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MMT

Trigger	Name the condition that activates MMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-11 - Moussaif Mechanism-Trace Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-11 - Moussaif Mechanism-Trace Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MMT transforms an admissible state into mechanism trace graph plus an uncertainty and provenance record. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-11 - Moussaif Mechanism-Trace Theory: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-11 - Moussaif Mechanism-Trace Theory: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid mechanism trace graph. Boundary analysis for TH-11 - Moussaif Mechanism-Trace Theory: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MMT: outcome المقترح تحت baseline وbudget يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit mechanism trace graph; 7) write telemetry and residual uncertainty. Integration analysis for TH-11 - Moussaif Mechanism-Trace Theory: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in trace completeness, reproducibility, diagnostic localization exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-11 - Moussaif Mechanism-Trace Theory: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MWSM, MAOLT, MEVE. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-11 - Moussaif Mechanism-Trace Theory: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MMT

Trigger	Name the condition that activates MMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-11 - Moussaif Mechanism-Trace Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-11 - Moussaif Mechanism-Trace Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MMT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-11 - Moussaif Mechanism-Trace Theory: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: mechanism trace graph; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-11 - Moussaif Mechanism-Trace Theory: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-11 - Moussaif Mechanism-Trace Theory: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MMT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-11 - Moussaif Mechanism-Trace Theory: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure trace completeness, reproducibility, diagnostic localization. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-11 - Moussaif Mechanism-Trace Theory: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by black-box success that cannot be explained, reproduced or repaired, the baseline produces an answer directly. The MMT route emits mechanism trace graph, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-11 - Moussaif Mechanism-Trace Theory: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MMT

Trigger	Name the condition that activates MMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-11 - Moussaif Mechanism-Trace Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-11 - Moussaif Mechanism-Trace Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where important outcomes require a trace from causal mechanism to evidence and world effect. Match models, prompts, budget and tools. If MMT does not improve trace completeness, reproducibility, diagnostic localization, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-11 - Moussaif Mechanism-Trace Theory: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: post-hoc stories mistaken for causal explanations. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-11 - Moussaif Mechanism-Trace Theory: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MMT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MWSM, MAOLT, MEVE), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-11 - Moussaif Mechanism-Trace Theory: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: trace completeness, reproducibility, diagnostic localization. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-11 - Moussaif Mechanism-Trace Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MMT

Trigger	Name the condition that activates MMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-11 - Moussaif Mechanism-Trace Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-12 - Moussaif World-State Model: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: plans based on stale, inferred or imagined external state. Core claim: observed, inferred, intended, simulated and unknown state must remain distinct. Mechanism specification for TH-12 - Moussaif World-State Model: translate research thesis into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

PRIMARY OPERATIONAL OBJECT

The required artifact is timestamped world-state snapshot. A verbal explanation without this object is not activation of MWSM. The object must be schema-valid, versioned and linked to the task contract. Integration analysis for TH-12 - Moussaif World-State Model: map primary operational object to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DEPENDENCIES

Principal dependencies and neighboring constructs: MPWG, MTPG, MAECT. Integration should preserve separate responsibilities and permit component-level ablation. Boundary analysis for TH-12 - Moussaif World-State Model: define inclusion, exclusion and abstention regions for dependencies. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

RESEARCH QUESTIONS

Can MWSM improve state accuracy, staleness, reconciliation delay under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Risk and counterexample analysis for TH-12 - Moussaif World-State Model: turn the named limitation in research questions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MWSM: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets plans based on stale, inferred or imagined external state. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Integration analysis for TH-12 - Moussaif World-State Model: map governing problem to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BOUNDARY AND EXCLUSIONS

MWSM does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Boundary analysis for TH-12 - Moussaif World-State Model: define inclusion, exclusion and abstention regions for boundary and exclusions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

KNOWN RISK

Primary conceptual risk: state explosion and false precision. The registry must record counterexamples, unresolved definitions and alternative explanations. Artifact discipline for TH-12 - Moussaif World-State Model: the known risk claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MWSM

Trigger	Name the condition that activates MWSM; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-12 - Moussaif World-State Model.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-12 - Moussaif World-State Model: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MWSM transforms an admissible state into timestamped world-state snapshot plus an uncertainty and provenance record. Artifact discipline for TH-12 - Moussaif World-State Model: the state and notation claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Integration analysis for TH-12 - Moussaif World-State Model: map invariants to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Research-program analysis for TH-12 - Moussaif World-State Model: decompose complexity model into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid timestamped world-state snapshot. Risk and counterexample analysis for TH-12 - Moussaif World-State Model: turn the named limitation in formal failure into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MWSM: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit timestamped world-state snapshot; 7) write telemetry and residual uncertainty. Research-program analysis for TH-12 - Moussaif World-State Model: decompose mechanism into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DECISION RULE

Adopt activation when expected gain in state accuracy, staleness, reconciliation delay exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Integration analysis for TH-12 - Moussaif World-State Model: map decision rule to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MPWG, MTPG, MAECT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Risk and counterexample analysis for TH-12 - Moussaif World-State Model: turn the named limitation in interfaces into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MWSM

Trigger	Name the condition that activates MWSM; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-12 - Moussaif World-State Model.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-12 - Moussaif World-State Model: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MWSM is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Mechanism specification for TH-12 - Moussaif World-State Model: translate runtime architecture into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OUTPUT CONTRACT

Output: timestamped world-state snapshot; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Mechanism specification for TH-12 - Moussaif World-State Model: translate output contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Research-program analysis for TH-12 - Moussaif World-State Model: decompose security and authority into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Artifact discipline for TH-12 - Moussaif World-State Model: the integration verdicts claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MWSM: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Risk and counterexample analysis for TH-12 - Moussaif World-State Model: turn the named limitation in input contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure state accuracy, staleness, reconciliation delay. Sensitive content is redacted or referenced, not copied into telemetry. Operational claim analysis for TH-12 - Moussaif World-State Model: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For observability, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

WORKED SCENARIO

Given a task dominated by plans based on stale, inferred or imagined external state, the baseline produces an answer directly. The MWSM route emits timestamped world-state snapshot, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Boundary analysis for TH-12 - Moussaif World-State Model: define inclusion, exclusion and abstention regions for worked scenario. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MWSM

Trigger	Name the condition that activates MWSM; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-12 - Moussaif World-State Model.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-12 - Moussaif World-State Model: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where observed, inferred, intended, simulated and unknown state must remain distinct. Match models, prompts, budget and tools. If MWSM does not improve state accuracy, staleness, reconciliation delay, the claimed mechanism is not supported under that regime. Artifact discipline for TH-12 - Moussaif World-State Model: the primary falsifier claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Integration analysis for TH-12 - Moussaif World-State Model: map baselines to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: state explosion and false precision. Artifact discipline for TH-12 - Moussaif World-State Model: the adversarial cases claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Risk and counterexample analysis for TH-12 - Moussaif World-State Model: turn the named limitation in next research action into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MWSM: outcome المقترح يصبح ذا قيمة فقط عندما يغير baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MPWG, MTPG, MAECT), the court and the memory write separately. Measure both capability loss and cost reduction. Risk and counterexample analysis for TH-12 - Moussaif World-State Model: turn the named limitation in ablation plan into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

METRICS

Primary metrics: state accuracy, staleness, reconciliation delay. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Artifact discipline for TH-12 - Moussaif World-State Model: the metrics claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Artifact discipline for TH-12 - Moussaif World-State Model: the evidence ladder claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MWSM

Trigger	Name the condition that activates MWSM; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-12 - Moussaif World-State Model.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-13 - Moussaif Project-World Graph: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: project plans detached from actors, assets, constraints and external dependencies. Core claim: a living graph should connect internal project structure to the world it can change. Operational claim analysis for TH-13 - Moussaif Project-World Graph: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research thesis, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

PRIMARY OPERATIONAL OBJECT

The required artifact is project-world graph. A verbal explanation without this object is not activation of MPWG. The object must be schema-valid, versioned and linked to the task contract. Boundary analysis for TH-13 - Moussaif Project-World Graph: define inclusion, exclusion and abstention regions for primary operational object. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DEPENDENCIES

Principal dependencies and neighboring constructs: MWSM, MTPG, MEVE. Integration should preserve separate responsibilities and permit component-level ablation. Artifact discipline for TH-13 - Moussaif Project-World Graph: the dependencies claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

RESEARCH QUESTIONS

Can MPWG improve dependency coverage, orphan rate, impact traceability under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Integration analysis for TH-13 - Moussaif Project-World Graph: map research questions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MPWG: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets project plans detached from actors, assets, constraints and external dependencies. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Boundary analysis for TH-13 - Moussaif Project-World Graph: define inclusion, exclusion and abstention regions for governing problem. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BOUNDARY AND EXCLUSIONS

MPWG does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Artifact discipline for TH-13 - Moussaif Project-World Graph: the boundary and exclusions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

KNOWN RISK

Primary conceptual risk: graph maintenance cost and outdated edges. The registry must record counterexamples, unresolved definitions and alternative explanations. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate known risk into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MPWG

Trigger	Name the condition that activates MPWG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-13 - Moussaif Project-World Graph.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-13 - Moussaif Project-World Graph: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MPWG transforms an admissible state into project-world graph plus an uncertainty and provenance record. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate state and notation into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Boundary analysis for TH-13 - Moussaif Project-World Graph: define inclusion, exclusion and abstention regions for invariants. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Risk and counterexample analysis for TH-13 - Moussaif Project-World Graph: turn the named limitation in complexity model into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid project-world graph. Integration analysis for TH-13 - Moussaif Project-World Graph: map formal failure to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MPWG: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit project-world graph; 7) write telemetry and residual uncertainty. Risk and counterexample analysis for TH-13 - Moussaif Project-World Graph: turn the named limitation in mechanism into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DECISION RULE

Adopt activation when expected gain in dependency coverage, orphan rate, impact traceability exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Boundary analysis for TH-13 - Moussaif Project-World Graph: define inclusion, exclusion and abstention regions for decision rule. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MWSM, MTPG, MEVE. Outputs feed planning, routing, courts, deployment or memory according to the object type. Integration analysis for TH-13 - Moussaif Project-World Graph: map interfaces to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MPWG

Trigger	Name the condition that activates MPWG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-13 - Moussaif Project-World Graph.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-13 - Moussaif Project-World Graph: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MPWG is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Operational claim analysis for TH-13 - Moussaif Project-World Graph: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For runtime architecture, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OUTPUT CONTRACT

Output: project-world graph; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Operational claim analysis for TH-13 - Moussaif Project-World Graph: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For output contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Risk and counterexample analysis for TH-13 - Moussaif Project-World Graph: turn the named limitation in security and authority into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate integration verdicts into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MPWG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Integration analysis for TH-13 - Moussaif Project-World Graph: map input contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure dependency coverage, orphan rate, impact traceability. Sensitive content is redacted or referenced, not copied into telemetry. Research-program analysis for TH-13 - Moussaif Project-World Graph: decompose observability into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

WORKED SCENARIO

Given a task dominated by project plans detached from actors, assets, constraints and external dependencies, the baseline produces an answer directly. The MPWG route emits project-world graph, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Artifact discipline for TH-13 - Moussaif Project-World Graph: the worked scenario claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MPWG

Trigger	Name the condition that activates MPWG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-13 - Moussaif Project-World Graph.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-13 - Moussaif Project-World Graph: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where a living graph should connect internal project structure to the world it can change. Match models, prompts, budget and tools. If MPWG does not improve dependency coverage, orphan rate, impact traceability, the claimed mechanism is not supported under that regime. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate primary falsifier into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Boundary analysis for TH-13 - Moussaif Project-World Graph: define inclusion, exclusion and abstention regions for baselines. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: graph maintenance cost and outdated edges. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate adversarial cases into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Integration analysis for TH-13 - Moussaif Project-World Graph: map next research action to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MPWG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MWSM, MTPG, MEVE), the court and the memory write separately. Measure both capability loss and cost reduction. Integration analysis for TH-13 - Moussaif Project-World Graph: map ablation plan to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

METRICS

Primary metrics: dependency coverage, orphan rate, impact traceability. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate metrics into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Mechanism specification for TH-13 - Moussaif Project-World Graph: translate evidence ladder into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MPWG

Trigger	Name the condition that activates MPWG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-13 - Moussaif Project-World Graph.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: knowledge stored as passive documents rather than executable operational memory. Core claim: knowledge units should expose triggers, dependencies, provenance, tests and actions. Operational claim analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research thesis, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

PRIMARY OPERATIONAL OBJECT

The required artifact is HyperKB card set and knowledge runtime. A verbal explanation without this object is not activation of MHKOS. The object must be schema-valid, versioned and linked to the task contract. Boundary analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: define inclusion, exclusion and abstention regions for primary operational object. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DEPENDENCIES

Principal dependencies and neighboring constructs: MPART, MCCT, MHSG. Integration should preserve separate responsibilities and permit component-level ablation. Artifact discipline for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: the dependencies claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

RESEARCH QUESTIONS

Can MHKOS improve retrieval precision, actionability, provenance completeness under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Integration analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: map research questions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MHKOS: outcome المقترح يصبح ذا قيمة فقط عندما يغير baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets knowledge stored as passive documents rather than executable operational memory. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Boundary analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: define inclusion, exclusion and abstention regions for governing problem. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BOUNDARY AND EXCLUSIONS

MHKOS does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Artifact discipline for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: the boundary and exclusions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

KNOWN RISK

Primary conceptual risk: over-structuring, retrieval brittleness and schema lock-in. The registry must record counterexamples, unresolved definitions and alternative explanations. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate known risk into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MHKOS

Trigger	Name the condition that activates MHKOS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-14 - Moussaif HyperKnowledge KB-as-OS Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MHKOS transforms an admissible state into HyperKB card set and knowledge runtime plus an uncertainty and provenance record. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate state and notation into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Boundary analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: define inclusion, exclusion and abstention regions for invariants. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Risk and counterexample analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: turn the named limitation in complexity model into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid HyperKB card set and knowledge runtime. Integration analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: map formal failure to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MHKOS: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit HyperKB card set and knowledge runtime; 7) write telemetry and residual uncertainty. Risk and counterexample analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: turn the named limitation in mechanism into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DECISION RULE

Adopt activation when expected gain in retrieval precision, actionability, provenance completeness exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Boundary analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: define inclusion, exclusion and abstention regions for decision rule. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MPART, MCCT, MHSG. Outputs feed planning, routing, courts, deployment or memory according to the object type. Integration analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: map interfaces to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MHKOS

Trigger	Name the condition that activates MHKOS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-14 - Moussaif HyperKnowledge KB-as-OS Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MHKOS is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Operational claim analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For runtime architecture, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OUTPUT CONTRACT

Output: HyperKB card set and knowledge runtime; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Operational claim analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For output contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Risk and counterexample analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: turn the named limitation in security and authority into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate integration verdicts into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MHKOS: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Integration analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: map input contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure retrieval precision, actionability, provenance completeness. Sensitive content is redacted or referenced, not copied into telemetry. Research-program analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: decompose observability into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

WORKED SCENARIO

Given a task dominated by knowledge stored as passive documents rather than executable operational memory, the baseline produces an answer directly. The MHKOS route emits HyperKB card set and knowledge runtime, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Artifact discipline for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: the worked scenario claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MHKOS

Trigger	Name the condition that activates MHKOS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-14 - Moussaif HyperKnowledge KB-as-OS Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where knowledge units should expose triggers, dependencies, provenance, tests and actions. Match models, prompts, budget and tools. If MHKOS does not improve retrieval precision, actionability, provenance completeness, the claimed mechanism is not supported under that regime. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate primary falsifier into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Boundary analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: define inclusion, exclusion and abstention regions for baselines. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: over-structuring, retrieval brittleness and schema lock-in. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate adversarial cases into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Integration analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: map next research action to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MHKOS: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MPART, MCCT, MHSG), the court and the memory write separately. Measure both capability loss and cost reduction. Integration analysis for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: map ablation plan to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

METRICS

Primary metrics: retrieval precision, actionability, provenance completeness. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate metrics into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Mechanism specification for TH-14 - Moussaif HyperKnowledge KB-as-OS Theory: translate evidence ladder into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MHKOS

Trigger	Name the condition that activates MHKOS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-14 - Moussaif HyperKnowledge KB-as-OS Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-15 - Moussaif Provenance-Aware Retrieval Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: retrieval that returns fluent but stale, duplicated or unattributed context. Core claim: ranking must include provenance, freshness, independence and task relevance. Boundary analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is provenance-filtered retrieval packet. A verbal explanation without this object is not activation of MPART. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: MELT, MHKOS, MCCT. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MPART improve precision, freshness, attribution, contradiction rate under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MPART: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets retrieval that returns fluent but stale, duplicated or unattributed context. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MPART does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: source popularity replacing source quality. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MPART

Trigger	Name the condition that activates MPART; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-15 - Moussaif Provenance-Aware Retrieval Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-15 - Moussaif Provenance-Aware Retrieval Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MPART transforms an admissible state into provenance-filtered retrieval packet plus an uncertainty and provenance record. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-15 - Moussaif Provenance-Aware Retrieval Theory: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid provenance-filtered retrieval packet. Operational claim analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MPART: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit provenance-filtered retrieval packet; 7) write telemetry and residual uncertainty. Mechanism specification for TH-15 - Moussaif Provenance-Aware Retrieval Theory: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in precision, freshness, attribution, contradiction rate exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MELT, MHKOS, MCCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MPART

Trigger	Name the condition that activates MPART; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-15 - Moussaif Provenance-Aware Retrieval Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-15 - Moussaif Provenance-Aware Retrieval Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MPART is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: provenance-filtered retrieval packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-15 - Moussaif Provenance-Aware Retrieval Theory: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MPART: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure precision, freshness, attribution, contradiction rate. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-15 - Moussaif Provenance-Aware Retrieval Theory: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by retrieval that returns fluent but stale, duplicated or unattributed context, the baseline produces an answer directly. The MPART route emits provenance-filtered retrieval packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MPART

Trigger	Name the condition that activates MPART; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-15 - Moussaif Provenance-Aware Retrieval Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-15 - Moussaif Provenance-Aware Retrieval Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where ranking must include provenance, freshness, independence and task relevance. Match models, prompts, budget and tools. If MPART does not improve precision, freshness, attribution, contradiction rate, the claimed mechanism is not supported under that regime. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: source popularity replacing source quality. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MPART: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MELT, MHKOS, MCCT), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: precision, freshness, attribution, contradiction rate. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-15 - Moussaif Provenance-Aware Retrieval Theory: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MPART

Trigger	Name the condition that activates MPART; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-15 - Moussaif Provenance-Aware Retrieval Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-16 - Moussaif Stratified Memory Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: one memory store mixing transient context, episodes, procedures and canonical facts. Core claim: memory strata need separate promotion, expiry, privacy and conflict rules. Research-program analysis for TH-16 - Moussaif Stratified Memory Theory: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is stratified memory write plan. A verbal explanation without this object is not activation of MSMT. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-16 - Moussaif Stratified Memory Theory: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: MPART, MCCT, MAOLT. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-16 - Moussaif Stratified Memory Theory: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MSMT improve usefulness, conflict rate, stale-memory rate, privacy incidents under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-16 - Moussaif Stratified Memory Theory: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MSMT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets one memory store mixing transient context, episodes, procedures and canonical facts. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-16 - Moussaif Stratified Memory Theory: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MSMT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-16 - Moussaif Stratified Memory Theory: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: premature canonization and irrecoverable deletion. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MSMT

Trigger	Name the condition that activates MSMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-16 - Moussaif Stratified Memory Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-16 - Moussaif Stratified Memory Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MSMT transforms an admissible state into stratified memory write plan plus an uncertainty and provenance record. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-16 - Moussaif Stratified Memory Theory: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-16 - Moussaif Stratified Memory Theory: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid stratified memory write plan. Boundary analysis for TH-16 - Moussaif Stratified Memory Theory: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MSMT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit stratified memory write plan; 7) write telemetry and residual uncertainty. Integration analysis for TH-16 - Moussaif Stratified Memory Theory: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in usefulness, conflict rate, stale-memory rate, privacy incidents exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-16 - Moussaif Stratified Memory Theory: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MPART, MCCT, MAOLT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-16 - Moussaif Stratified Memory Theory: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MSMT

Trigger	Name the condition that activates MSMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-16 - Moussaif Stratified Memory Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-16 - Moussaif Stratified Memory Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MSMT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-16 - Moussaif Stratified Memory Theory: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: stratified memory write plan; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-16 - Moussaif Stratified Memory Theory: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-16 - Moussaif Stratified Memory Theory: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MSMT: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-16 - Moussaif Stratified Memory Theory: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure usefulness, conflict rate, stale-memory rate, privacy incidents. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-16 - Moussaif Stratified Memory Theory: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by one memory store mixing transient context, episodes, procedures and canonical facts, the baseline produces an answer directly. The MSMT route emits stratified memory write plan, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-16 - Moussaif Stratified Memory Theory: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MSMT

Trigger	Name the condition that activates MSMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-16 - Moussaif Stratified Memory Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-16 - Moussaif Stratified Memory Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where memory strata need separate promotion, expiry, privacy and conflict rules. Match models, prompts, budget and tools. If MSMT does not improve usefulness, conflict rate, stale-memory rate, privacy incidents, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-16 - Moussaif Stratified Memory Theory: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: premature canonization and irrecoverable deletion. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-16 - Moussaif Stratified Memory Theory: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MSMT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MPART, MCCT, MAOLT), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-16 - Moussaif Stratified Memory Theory: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: usefulness, conflict rate, stale-memory rate, privacy incidents. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-16 - Moussaif Stratified Memory Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MSMT

Trigger	Name the condition that activates MSMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-16 - Moussaif Stratified Memory Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-17 - Moussaif Context-Capsule Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: long contexts that are costly, noisy and hard to hand off. Core claim: a bounded capsule can preserve governing intent, evidence, state and unresolved items. Artifact discipline for TH-17 - Moussaif Context-Capsule Theory: the research thesis claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

PRIMARY OPERATIONAL OBJECT

The required artifact is context capsule. A verbal explanation without this object is not activation of MCCT. The object must be schema-valid, versioned and linked to the task contract. Risk and counterexample analysis for TH-17 - Moussaif Context-Capsule Theory: turn the named limitation in primary operational object into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DEPENDENCIES

Principal dependencies and neighboring constructs: MGCG, MSMT, MHSO. Integration should preserve separate responsibilities and permit component-level ablation. Integration analysis for TH-17 - Moussaif Context-Capsule Theory: map dependencies to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

RESEARCH QUESTIONS

Can MCCT improve compression fidelity, handoff success, token reduction under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Research-program analysis for TH-17 - Moussaif Context-Capsule Theory: decompose research questions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MCCT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets long contexts that are costly, noisy and hard to hand off. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Risk and counterexample analysis for TH-17 - Moussaif Context-Capsule Theory: turn the named limitation in governing problem into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BOUNDARY AND EXCLUSIONS

MCCT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Integration analysis for TH-17 - Moussaif Context-Capsule Theory: map boundary and exclusions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

KNOWN RISK

Primary conceptual risk: compression loss and hidden assumptions. The registry must record counterexamples, unresolved definitions and alternative explanations. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for known risk. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MCCT

Trigger	Name the condition that activates MCCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-17 - Moussaif Context-Capsule Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-17 - Moussaif Context-Capsule Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MCCT transforms an admissible state into context capsule plus an uncertainty and provenance record. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for state and notation. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Risk and counterexample analysis for TH-17 - Moussaif Context-Capsule Theory: turn the named limitation in invariants into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Operational claim analysis for TH-17 - Moussaif Context-Capsule Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For complexity model, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid context capsule. Research-program analysis for TH-17 - Moussaif Context-Capsule Theory: decompose formal failure into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MCCT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit context capsule; 7) write telemetry and residual uncertainty. Operational claim analysis for TH-17 - Moussaif Context-Capsule Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For mechanism, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DECISION RULE

Adopt activation when expected gain in compression fidelity, handoff success, token reduction exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Risk and counterexample analysis for TH-17 - Moussaif Context-Capsule Theory: turn the named limitation in decision rule into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MGCG, MSMT, MHSO. Outputs feed planning, routing, courts, deployment or memory according to the object type. Research-program analysis for TH-17 - Moussaif Context-Capsule Theory: decompose interfaces into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MCCT

Trigger	Name the condition that activates MCCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-17 - Moussaif Context-Capsule Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-17 - Moussaif Context-Capsule Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MCCT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Artifact discipline for TH-17 - Moussaif Context-Capsule Theory: the runtime architecture claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OUTPUT CONTRACT

Output: context capsule; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Artifact discipline for TH-17 - Moussaif Context-Capsule Theory: the output contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Operational claim analysis for TH-17 - Moussaif Context-Capsule Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For security and authority, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for integration verdicts. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MCCT: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Research-program analysis for TH-17 - Moussaif Context-Capsule Theory: decompose input contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure compression fidelity, handoff success, token reduction. Sensitive content is redacted or referenced, not copied into telemetry. Mechanism specification for TH-17 - Moussaif Context-Capsule Theory: translate observability into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

WORKED SCENARIO

Given a task dominated by long contexts that are costly, noisy and hard to hand off, the baseline produces an answer directly. The MCCT route emits context capsule, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Integration analysis for TH-17 - Moussaif Context-Capsule Theory: map worked scenario to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCCT

Trigger	Name the condition that activates MCCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-17 - Moussaif Context-Capsule Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-17 - Moussaif Context-Capsule Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where a bounded capsule can preserve governing intent, evidence, state and unresolved items. Match models, prompts, budget and tools. If MCCT does not improve compression fidelity, handoff success, token reduction, the claimed mechanism is not supported under that regime. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for primary falsifier. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Risk and counterexample analysis for TH-17 - Moussaif Context-Capsule Theory: turn the named limitation in baselines into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: compression loss and hidden assumptions. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for adversarial cases. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Research-program analysis for TH-17 - Moussaif Context-Capsule Theory: decompose next research action into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MCCT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MGCG, MSMT, MHSO), the court and the memory write separately. Measure both capability loss and cost reduction. Research-program analysis for TH-17 - Moussaif Context-Capsule Theory: decompose ablation plan into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

METRICS

Primary metrics: compression fidelity, handoff success, token reduction. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for metrics. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Boundary analysis for TH-17 - Moussaif Context-Capsule Theory: define inclusion, exclusion and abstention regions for evidence ladder. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MCCT

Trigger	Name the condition that activates MCCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-17 - Moussaif Context-Capsule Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-18 - Moussaif HyperSkill-Graph Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: flat skill libraries with duplication and unclear composition. Core claim: skills should form a typed graph of triggers, prerequisites, outputs and tests. Integration analysis for TH-18 - Moussaif HyperSkill-Graph Theory: map research thesis to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

PRIMARY OPERATIONAL OBJECT

The required artifact is HyperSkill graph. A verbal explanation without this object is not activation of MHSG. The object must be schema-valid, versioned and linked to the task contract. Operational claim analysis for TH-18 - Moussaif HyperSkill-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary operational object, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DEPENDENCIES

Principal dependencies and neighboring constructs: MHKOS, MSSDT, MSHR. Integration should preserve separate responsibilities and permit component-level ablation. Research-program analysis for TH-18 - Moussaif HyperSkill-Graph Theory: decompose dependencies into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

RESEARCH QUESTIONS

Can MHSG improve coverage, reuse, activation precision, duplication under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Mechanism specification for TH-18 - Moussaif HyperSkill-Graph Theory: translate research questions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MHSG: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets flat skill libraries with duplication and unclear composition. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Operational claim analysis for TH-18 - Moussaif HyperSkill-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For governing problem, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BOUNDARY AND EXCLUSIONS

MHSG does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Research-program analysis for TH-18 - Moussaif HyperSkill-Graph Theory: decompose boundary and exclusions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

KNOWN RISK

Primary conceptual risk: taxonomy bloat and cyclic dependencies. The registry must record counterexamples, unresolved definitions and alternative explanations. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in known risk into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MHSG

Trigger	Name the condition that activates MHSG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-18 - Moussaif HyperSkill-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-18 - Moussaif HyperSkill-Graph Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MHSG transforms an admissible state into HyperSkill graph plus an uncertainty and provenance record. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in state and notation into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Operational claim analysis for TH-18 - Moussaif HyperSkill-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For invariants, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Artifact discipline for TH-18 - Moussaif HyperSkill-Graph Theory: the complexity model claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid HyperSkill graph. Mechanism specification for TH-18 - Moussaif HyperSkill-Graph Theory: translate formal failure into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MHSG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit HyperSkill graph; 7) write telemetry and residual uncertainty. Artifact discipline for TH-18 - Moussaif HyperSkill-Graph Theory: the mechanism claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DECISION RULE

Adopt activation when expected gain in coverage, reuse, activation precision, duplication exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Operational claim analysis for TH-18 - Moussaif HyperSkill-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For decision rule, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MHKOS, MSSDT, MSHR. Outputs feed planning, routing, courts, deployment or memory according to the object type. Mechanism specification for TH-18 - Moussaif HyperSkill-Graph Theory: translate interfaces into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MHSG

Trigger	Name the condition that activates MHSG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-18 - Moussaif HyperSkill-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-18 - Moussaif HyperSkill-Graph Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MHSG is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Integration analysis for TH-18 - Moussaif HyperSkill-Graph Theory: map runtime architecture to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OUTPUT CONTRACT

Output: HyperSkill graph; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Integration analysis for TH-18 - Moussaif HyperSkill-Graph Theory: map output contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Artifact discipline for TH-18 - Moussaif HyperSkill-Graph Theory: the security and authority claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in integration verdicts into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MHSG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Mechanism specification for TH-18 - Moussaif HyperSkill-Graph Theory: translate input contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure coverage, reuse, activation precision, duplication. Sensitive content is redacted or referenced, not copied into telemetry. Boundary analysis for TH-18 - Moussaif HyperSkill-Graph Theory: define inclusion, exclusion and abstention regions for observability. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

WORKED SCENARIO

Given a task dominated by flat skill libraries with duplication and unclear composition, the baseline produces an answer directly. The MHSG route emits HyperSkill graph, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Research-program analysis for TH-18 - Moussaif HyperSkill-Graph Theory: decompose worked scenario into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MHSG

Trigger	Name the condition that activates MHSG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-18 - Moussaif HyperSkill-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-18 - Moussaif HyperSkill-Graph Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where skills should form a typed graph of triggers, prerequisites, outputs and tests. Match models, prompts, budget and tools. If MHSG does not improve coverage, reuse, activation precision, duplication, the claimed mechanism is not supported under that regime. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in primary falsifier into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Operational claim analysis for TH-18 - Moussaif HyperSkill-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For baselines, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: taxonomy bloat and cyclic dependencies. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in adversarial cases into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Mechanism specification for TH-18 - Moussaif HyperSkill-Graph Theory: translate next research action into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MHSG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MHKOS, MSSDT, MSHR), the court and the memory write separately. Measure both capability loss and cost reduction. Mechanism specification for TH-18 - Moussaif HyperSkill-Graph Theory: translate ablation plan into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

METRICS

Primary metrics: coverage, reuse, activation precision, duplication. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in metrics into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Risk and counterexample analysis for TH-18 - Moussaif HyperSkill-Graph Theory: turn the named limitation in evidence ladder into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MHSG

Trigger	Name the condition that activates MHSG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-18 - Moussaif HyperSkill-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-19 - Moussaif Sparse HyperRouting Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: loading too many skills, models and knowledge units for every task. Core claim: select the smallest sufficient capability set and escalate only on evidence. Research-program analysis for TH-19 - Moussaif Sparse HyperRouting Theory: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is sparse route packet. A verbal explanation without this object is not activation of MSHR. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-19 - Moussaif Sparse HyperRouting Theory: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: MHSG, MMPR, MCVT. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-19 - Moussaif Sparse HyperRouting Theory: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MSHR improve quality per cost, route sparsity, escalation accuracy under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-19 - Moussaif Sparse HyperRouting Theory: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MSHR: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets loading too many skills, models and knowledge units for every task. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-19 - Moussaif Sparse HyperRouting Theory: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MSHR does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-19 - Moussaif Sparse HyperRouting Theory: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: starvation of weak signals and brittle router confidence. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MSHR

Trigger	Name the condition that activates MSHR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-19 - Moussaif Sparse HyperRouting Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-19 - Moussaif Sparse HyperRouting Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MSHR transforms an admissible state into sparse route packet plus an uncertainty and provenance record. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-19 - Moussaif Sparse HyperRouting Theory: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-19 - Moussaif Sparse HyperRouting Theory: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid sparse route packet. Boundary analysis for TH-19 - Moussaif Sparse HyperRouting Theory: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MSHR: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit sparse route packet; 7) write telemetry and residual uncertainty. Integration analysis for TH-19 - Moussaif Sparse HyperRouting Theory: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in quality per cost, route sparsity, escalation accuracy exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-19 - Moussaif Sparse HyperRouting Theory: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MHSG, MMPR, MCVT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-19 - Moussaif Sparse HyperRouting Theory: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MSHR

Trigger	Name the condition that activates MSHR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-19 - Moussaif Sparse HyperRouting Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-19 - Moussaif Sparse HyperRouting Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MSHR is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-19 - Moussaif Sparse HyperRouting Theory: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: sparse route packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-19 - Moussaif Sparse HyperRouting Theory: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-19 - Moussaif Sparse HyperRouting Theory: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MSHR: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-19 - Moussaif Sparse HyperRouting Theory: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure quality per cost, route sparsity, escalation accuracy. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-19 - Moussaif Sparse HyperRouting Theory: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by loading too many skills, models and knowledge units for every task, the baseline produces an answer directly. The MSHR route emits sparse route packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-19 - Moussaif Sparse HyperRouting Theory: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MSHR

Trigger	Name the condition that activates MSHR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-19 - Moussaif Sparse HyperRouting Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-19 - Moussaif Sparse HyperRouting Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where select the smallest sufficient capability set and escalate only on evidence. Match models, prompts, budget and tools. If MSHR does not improve quality per cost, route sparsity, escalation accuracy, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-19 - Moussaif Sparse HyperRouting Theory: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: starvation of weak signals and brittle router confidence. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-19 - Moussaif Sparse HyperRouting Theory: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MSHR: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MHSG, MMPR, MCVT), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-19 - Moussaif Sparse HyperRouting Theory: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: quality per cost, route sparsity, escalation accuracy. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-19 - Moussaif Sparse HyperRouting Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MSHR

Trigger	Name the condition that activates MSHR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-19 - Moussaif Sparse HyperRouting Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-20 - Moussaif Bounded HyperLoop Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: iterative agents that loop without measurable progress or safe stopping. Core claim: every loop requires an evidence-gain objective, budget and stop predicate. Research-program analysis for TH-20 - Moussaif Bounded HyperLoop Theory: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is bounded loop contract. A verbal explanation without this object is not activation of MBHL. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-20 - Moussaif Bounded HyperLoop Theory: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: MGCG, MAECT, MDBCT. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-20 - Moussaif Bounded HyperLoop Theory: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MBHL improve evidence gain, convergence, cost, timeout rate under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-20 - Moussaif Bounded HyperLoop Theory: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MBHL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets iterative agents that loop without measurable progress or safe stopping. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-20 - Moussaif Bounded HyperLoop Theory: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MBHL does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-20 - Moussaif Bounded HyperLoop Theory: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: premature stopping or endless reformulation. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MBHL

Trigger	Name the condition that activates MBHL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-20 - Moussaif Bounded HyperLoop Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-20 - Moussaif Bounded HyperLoop Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MBHL transforms an admissible state into bounded loop contract plus an uncertainty and provenance record. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-20 - Moussaif Bounded HyperLoop Theory: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-20 - Moussaif Bounded HyperLoop Theory: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid bounded loop contract. Boundary analysis for TH-20 - Moussaif Bounded HyperLoop Theory: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MBHL: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit bounded loop contract; 7) write telemetry and residual uncertainty. Integration analysis for TH-20 - Moussaif Bounded HyperLoop Theory: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in evidence gain, convergence, cost, timeout rate exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-20 - Moussaif Bounded HyperLoop Theory: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MGCG, MAECT, MDBCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-20 - Moussaif Bounded HyperLoop Theory: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MBHL

Trigger	Name the condition that activates MBHL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-20 - Moussaif Bounded HyperLoop Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-20 - Moussaif Bounded HyperLoop Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MBHL is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-20 - Moussaif Bounded HyperLoop Theory: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: bounded loop contract; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-20 - Moussaif Bounded HyperLoop Theory: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-20 - Moussaif Bounded HyperLoop Theory: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MBHL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-20 - Moussaif Bounded HyperLoop Theory: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure evidence gain, convergence, cost, timeout rate. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-20 - Moussaif Bounded HyperLoop Theory: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by iterative agents that loop without measurable progress or safe stopping, the baseline produces an answer directly. The MBHL route emits bounded loop contract, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-20 - Moussaif Bounded HyperLoop Theory: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MBHL

Trigger	Name the condition that activates MBHL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-20 - Moussaif Bounded HyperLoop Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-20 - Moussaif Bounded HyperLoop Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where every loop requires an evidence-gain objective, budget and stop predicate. Match models, prompts, budget and tools. If MBHL does not improve evidence gain, convergence, cost, timeout rate, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-20 - Moussaif Bounded HyperLoop Theory: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: premature stopping or endless reformulation. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-20 - Moussaif Bounded HyperLoop Theory: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MBHL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MGCG, MAECT, MDBCT), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-20 - Moussaif Bounded HyperLoop Theory: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: evidence gain, convergence, cost, timeout rate. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-20 - Moussaif Bounded HyperLoop Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MBHL

Trigger	Name the condition that activates MBHL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-20 - Moussaif Bounded HyperLoop Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-21 - Moussaif Tool-Intelligence Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: tools treated as untyped extensions of language generation. Core claim: intelligent tool use requires affordance models, typed contracts and state reconciliation. Research-program analysis for TH-21 - Moussaif Tool-Intelligence Theory: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is tool contract and postcondition report. A verbal explanation without this object is not activation of MTIT. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-21 - Moussaif Tool-Intelligence Theory: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: MABT, MIRTT, MAOLT. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-21 - Moussaif Tool-Intelligence Theory: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MTIT improve success rate, misuse rate, state accuracy, rollback success under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-21 - Moussaif Tool-Intelligence Theory: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MTIT: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets tools treated as untyped extensions of language generation. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-21 - Moussaif Tool-Intelligence Theory: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MTIT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-21 - Moussaif Tool-Intelligence Theory: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: tool hallucination and false success inference. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MTIT

Trigger	Name the condition that activates MTIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-21 - Moussaif Tool-Intelligence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-21 - Moussaif Tool-Intelligence Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MTIT transforms an admissible state into tool contract and postcondition report plus an uncertainty and provenance record. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-21 - Moussaif Tool-Intelligence Theory: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-21 - Moussaif Tool-Intelligence Theory: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid tool contract and postcondition report. Boundary analysis for TH-21 - Moussaif Tool-Intelligence Theory: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MTIT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit tool contract and postcondition report; 7) write telemetry and residual uncertainty. Integration analysis for TH-21 - Moussaif Tool-Intelligence Theory: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in success rate, misuse rate, state accuracy, rollback success exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-21 - Moussaif Tool-Intelligence Theory: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MABT, MIRTT, MAOLT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-21 - Moussaif Tool-Intelligence Theory: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MTIT

Trigger	Name the condition that activates MTIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-21 - Moussaif Tool-Intelligence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-21 - Moussaif Tool-Intelligence Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MTIT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-21 - Moussaif Tool-Intelligence Theory: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: tool contract and postcondition report; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-21 - Moussaif Tool-Intelligence Theory: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-21 - Moussaif Tool-Intelligence Theory: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MTIT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-21 - Moussaif Tool-Intelligence Theory: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure success rate, misuse rate, state accuracy, rollback success. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-21 - Moussaif Tool-Intelligence Theory: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by tools treated as untyped extensions of language generation, the baseline produces an answer directly. The MTIT route emits tool contract and postcondition report, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-21 - Moussaif Tool-Intelligence Theory: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MTIT

Trigger	Name the condition that activates MTIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-21 - Moussaif Tool-Intelligence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-21 - Moussaif Tool-Intelligence Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where intelligent tool use requires affordance models, typed contracts and state reconciliation. Match models, prompts, budget and tools. If MTIT does not improve success rate, misuse rate, state accuracy, rollback success, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-21 - Moussaif Tool-Intelligence Theory: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: tool hallucination and false success inference. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-21 - Moussaif Tool-Intelligence Theory: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MTIT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MABT, MIRTT, MAOLT), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-21 - Moussaif Tool-Intelligence Theory: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: success rate, misuse rate, state accuracy, rollback success. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-21 - Moussaif Tool-Intelligence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MTIT

Trigger	Name the condition that activates MTIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-21 - Moussaif Tool-Intelligence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-22 - Moussaif Temporal Plan-Graph Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: linear plans unable to represent dependencies, timing and reversible branches. Core claim: plans should be temporal graphs with checkpoints, conditions and recovery edges. Operational claim analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research thesis, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

PRIMARY OPERATIONAL OBJECT

The required artifact is temporal plan graph. A verbal explanation without this object is not activation of MTPG. The object must be schema-valid, versioned and linked to the task contract. Boundary analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: define inclusion, exclusion and abstention regions for primary operational object. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DEPENDENCIES

Principal dependencies and neighboring constructs: MGCG, MWSM, MAECT. Integration should preserve separate responsibilities and permit component-level ablation. Artifact discipline for TH-22 - Moussaif Temporal Plan-Graph Theory: the dependencies claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

RESEARCH QUESTIONS

Can MTPG improve critical-path accuracy, replan rate, deadline adherence under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Integration analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: map research questions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MTPG: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets linear plans unable to represent dependencies, timing and reversible branches. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Boundary analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: define inclusion, exclusion and abstention regions for governing problem. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BOUNDARY AND EXCLUSIONS

MTPG does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Artifact discipline for TH-22 - Moussaif Temporal Plan-Graph Theory: the boundary and exclusions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

KNOWN RISK

Primary conceptual risk: graph overhead for trivial tasks and stale schedules. The registry must record counterexamples, unresolved definitions and alternative explanations. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate known risk into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MTPG

Trigger	Name the condition that activates MTPG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-22 - Moussaif Temporal Plan-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-22 - Moussaif Temporal Plan-Graph Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MTPG transforms an admissible state into temporal plan graph plus an uncertainty and provenance record. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate state and notation into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Boundary analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: define inclusion, exclusion and abstention regions for invariants. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Risk and counterexample analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: turn the named limitation in complexity model into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid temporal plan graph. Integration analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: map formal failure to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MTPG: outcome المقترح تحت baseline وbudget يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit temporal plan graph; 7) write telemetry and residual uncertainty. Risk and counterexample analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: turn the named limitation in mechanism into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DECISION RULE

Adopt activation when expected gain in critical-path accuracy, replan rate, deadline adherence exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Boundary analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: define inclusion, exclusion and abstention regions for decision rule. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MGCG, MWSM, MAECT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Integration analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: map interfaces to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MTPG

Trigger	Name the condition that activates MTPG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-22 - Moussaif Temporal Plan-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-22 - Moussaif Temporal Plan-Graph Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MTPG is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Operational claim analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For runtime architecture, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OUTPUT CONTRACT

Output: temporal plan graph; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Operational claim analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For output contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Risk and counterexample analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: turn the named limitation in security and authority into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate integration verdicts into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MTPG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Integration analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: map input contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure critical-path accuracy, replan rate, deadline adherence. Sensitive content is redacted or referenced, not copied into telemetry. Research-program analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: decompose observability into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

WORKED SCENARIO

Given a task dominated by linear plans unable to represent dependencies, timing and reversible branches, the baseline produces an answer directly. The MTPG route emits temporal plan graph, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Artifact discipline for TH-22 - Moussaif Temporal Plan-Graph Theory: the worked scenario claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MTPG

Trigger	Name the condition that activates MTPG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-22 - Moussaif Temporal Plan-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-22 - Moussaif Temporal Plan-Graph Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where plans should be temporal graphs with checkpoints, conditions and recovery edges. Match models, prompts, budget and tools. If MTPG does not improve critical-path accuracy, replan rate, deadline adherence, the claimed mechanism is not supported under that regime. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate primary falsifier into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Boundary analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: define inclusion, exclusion and abstention regions for baselines. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: graph overhead for trivial tasks and stale schedules. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate adversarial cases into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Integration analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: map next research action to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MTPG: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MGCG, MWSM, MAECT), the court and the memory write separately. Measure both capability loss and cost reduction. Integration analysis for TH-22 - Moussaif Temporal Plan-Graph Theory: map ablation plan to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

METRICS

Primary metrics: critical-path accuracy, replan rate, deadline adherence. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate metrics into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Mechanism specification for TH-22 - Moussaif Temporal Plan-Graph Theory: translate evidence ladder into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MTPG

Trigger	Name the condition that activates MTPG; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-22 - Moussaif Temporal Plan-Graph Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-23 - Moussaif Adaptive Execution-Control Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: plans that continue unchanged after evidence or world state shifts. Core claim: execution control should re-evaluate route, risk and stop conditions after state change. Mechanism specification for TH-23 - Moussaif Adaptive Execution-Control Theory: translate research thesis into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

PRIMARY OPERATIONAL OBJECT

The required artifact is execution control packet. A verbal explanation without this object is not activation of MAECT. The object must be schema-valid, versioned and linked to the task contract. Integration analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: map primary operational object to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DEPENDENCIES

Principal dependencies and neighboring constructs: MTPG, MWSM, MBHL. Integration should preserve separate responsibilities and permit component-level ablation. Boundary analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: define inclusion, exclusion and abstention regions for dependencies. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

RESEARCH QUESTIONS

Can MAECT improve adaptation latency, avoided failure, control stability under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Risk and counterexample analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: turn the named limitation in research questions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MAECT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets plans that continue unchanged after evidence or world state shifts. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Integration analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: map governing problem to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BOUNDARY AND EXCLUSIONS

MAECT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Boundary analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: define inclusion, exclusion and abstention regions for boundary and exclusions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

KNOWN RISK

Primary conceptual risk: oscillation and excessive replanning. The registry must record counterexamples, unresolved definitions and alternative explanations. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the known risk claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MAECT

Trigger	Name the condition that activates MAECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-23 - Moussaif Adaptive Execution-Control Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-23 - Moussaif Adaptive Execution-Control Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MAECT transforms an admissible state into execution control packet plus an uncertainty and provenance record. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the state and notation claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Integration analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: map invariants to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Research-program analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: decompose complexity model into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid execution control packet. Risk and counterexample analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: turn the named limitation in formal failure into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MAECT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit execution control packet; 7) write telemetry and residual uncertainty. Research-program analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: decompose mechanism into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DECISION RULE

Adopt activation when expected gain in adaptation latency, avoided failure, control stability exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Integration analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: map decision rule to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MTPG, MWSM, MBHL. Outputs feed planning, routing, courts, deployment or memory according to the object type. Risk and counterexample analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: turn the named limitation in interfaces into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MAECT

Trigger	Name the condition that activates MAECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-23 - Moussaif Adaptive Execution-Control Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-23 - Moussaif Adaptive Execution-Control Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MAECT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Mechanism specification for TH-23 - Moussaif Adaptive Execution-Control Theory: translate runtime architecture into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OUTPUT CONTRACT

Output: execution control packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Mechanism specification for TH-23 - Moussaif Adaptive Execution-Control Theory: translate output contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Research-program analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: decompose security and authority into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the integration verdicts claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MAECT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Risk and counterexample analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: turn the named limitation in input contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure adaptation latency, avoided failure, control stability. Sensitive content is redacted or referenced, not copied into telemetry. Operational claim analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For observability, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

WORKED SCENARIO

Given a task dominated by plans that continue unchanged after evidence or world state shifts, the baseline produces an answer directly. The MAECT route emits execution control packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Boundary analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: define inclusion, exclusion and abstention regions for worked scenario. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MAECT

Trigger	Name the condition that activates MAECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-23 - Moussaif Adaptive Execution-Control Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-23 - Moussaif Adaptive Execution-Control Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where execution control should re-evaluate route, risk and stop conditions after state change. Match models, prompts, budget and tools. If MAECT does not improve adaptation latency, avoided failure, control stability, the claimed mechanism is not supported under that regime. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the primary falsifier claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Integration analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: map baselines to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: oscillation and excessive replanning. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the adversarial cases claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Risk and counterexample analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: turn the named limitation in next research action into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MAECT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MTPG, MWSM, MBHL), the court and the memory write separately. Measure both capability loss and cost reduction. Risk and counterexample analysis for TH-23 - Moussaif Adaptive Execution-Control Theory: turn the named limitation in ablation plan into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

METRICS

Primary metrics: adaptation latency, avoided failure, control stability. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the metrics claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Artifact discipline for TH-23 - Moussaif Adaptive Execution-Control Theory: the evidence ladder claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MAECT

Trigger	Name the condition that activates MAECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-23 - Moussaif Adaptive Execution-Control Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-24 - Moussaif HyperSubagent Orchestration Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: multi-agent work with unclear contracts, handoffs and merge authority. Core claim: subagents require bounded scopes, typed handoffs and explicit merge rules. Artifact discipline for TH-24 - Moussaif HyperSubagent Orchestration Theory: the research thesis claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

PRIMARY OPERATIONAL OBJECT

The required artifact is subagent mission set and merge trace. A verbal explanation without this object is not activation of MHSO. The object must be schema-valid, versioned and linked to the task contract. Risk and counterexample analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: turn the named limitation in primary operational object into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DEPENDENCIES

Principal dependencies and neighboring constructs: MCCT, MTPG, MCIP. Integration should preserve separate responsibilities and permit component-level ablation. Integration analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: map dependencies to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

RESEARCH QUESTIONS

Can MHSO improve coordination loss, parallel speedup, merge defects under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Research-program analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: decompose research questions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MHSO: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets multi-agent work with unclear contracts, handoffs and merge authority. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Risk and counterexample analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: turn the named limitation in governing problem into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BOUNDARY AND EXCLUSIONS

MHSO does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Integration analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: map boundary and exclusions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

KNOWN RISK

Primary conceptual risk: agent proliferation and duplicated work. The registry must record counterexamples, unresolved definitions and alternative explanations. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for known risk. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MHSO

Trigger	Name the condition that activates MHSO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-24 - Moussaif HyperSubagent Orchestration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-24 - Moussaif HyperSubagent Orchestration Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MHSO transforms an admissible state into subagent mission set and merge trace plus an uncertainty and provenance record. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for state and notation. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Risk and counterexample analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: turn the named limitation in invariants into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Operational claim analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For complexity model, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid subagent mission set and merge trace. Research-program analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: decompose formal failure into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MHSO: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit subagent mission set and merge trace; 7) write telemetry and residual uncertainty. Operational claim analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For mechanism, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DECISION RULE

Adopt activation when expected gain in coordination loss, parallel speedup, merge defects exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Risk and counterexample analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: turn the named limitation in decision rule into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MCCT, MTPG, MCIP. Outputs feed planning, routing, courts, deployment or memory according to the object type. Research-program analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: decompose interfaces into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MHSO

Trigger	Name the condition that activates MHSO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-24 - Moussaif HyperSubagent Orchestration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-24 - Moussaif HyperSubagent Orchestration Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MHSO is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Artifact discipline for TH-24 - Moussaif HyperSubagent Orchestration Theory: the runtime architecture claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OUTPUT CONTRACT

Output: subagent mission set and merge trace; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Artifact discipline for TH-24 - Moussaif HyperSubagent Orchestration Theory: the output contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Operational claim analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For security and authority, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for integration verdicts. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MHSO: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Research-program analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: decompose input contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure coordination loss, parallel speedup, merge defects. Sensitive content is redacted or referenced, not copied into telemetry. Mechanism specification for TH-24 - Moussaif HyperSubagent Orchestration Theory: translate observability into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

WORKED SCENARIO

Given a task dominated by multi-agent work with unclear contracts, handoffs and merge authority, the baseline produces an answer directly. The MHSO route emits subagent mission set and merge trace, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Integration analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: map worked scenario to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MHSO

Trigger	Name the condition that activates MHSO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-24 - Moussaif HyperSubagent Orchestration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-24 - Moussaif HyperSubagent Orchestration Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where subagents require bounded scopes, typed handoffs and explicit merge rules. Match models, prompts, budget and tools. If MHSO does not improve coordination loss, parallel speedup, merge defects, the claimed mechanism is not supported under that regime. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for primary falsifier. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Risk and counterexample analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: turn the named limitation in baselines into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: agent proliferation and duplicated work. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for adversarial cases. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Research-program analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: decompose next research action into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MHSO: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MCCT, MTPG, MCIP), the court and the memory write separately. Measure both capability loss and cost reduction. Research-program analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: decompose ablation plan into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

METRICS

Primary metrics: coordination loss, parallel speedup, merge defects. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for metrics. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Boundary analysis for TH-24 - Moussaif HyperSubagent Orchestration Theory: define inclusion, exclusion and abstention regions for evidence ladder. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MHSO

Trigger	Name the condition that activates MHSO; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-24 - Moussaif HyperSubagent Orchestration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-25 - Moussaif Complementarity and Integration Protocol: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: components combined because they are individually impressive rather than jointly useful. Core claim: integration is justified only when components add non-redundant causal value. Boundary analysis for TH-25 - Moussaif Complementarity and Integration Protocol: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is complementarity matrix. A verbal explanation without this object is not activation of MCIP. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-25 - Moussaif Complementarity and Integration Protocol: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: MASO, MCD, MSSDT. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-25 - Moussaif Complementarity and Integration Protocol: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MCIP improve marginal capability gain, redundancy, interface defects under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-25 - Moussaif Complementarity and Integration Protocol: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCIP: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets components combined because they are individually impressive rather than jointly useful. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-25 - Moussaif Complementarity and Integration Protocol: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MCIP does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-25 - Moussaif Complementarity and Integration Protocol: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: false synergy and interface complexity. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCIP

Trigger	Name the condition that activates MCIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-25 - Moussaif Complementarity and Integration Protocol.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-25 - Moussaif Complementarity and Integration Protocol: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MCIP transforms an admissible state into complementarity matrix plus an uncertainty and provenance record. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-25 - Moussaif Complementarity and Integration Protocol: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-25 - Moussaif Complementarity and Integration Protocol: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid complementarity matrix. Operational claim analysis for TH-25 - Moussaif Complementarity and Integration Protocol: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCIP: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit complementarity matrix; 7) write telemetry and residual uncertainty. Mechanism specification for TH-25 - Moussaif Complementarity and Integration Protocol: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in marginal capability gain, redundancy, interface defects exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-25 - Moussaif Complementarity and Integration Protocol: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MASO, MCD, MSSDT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-25 - Moussaif Complementarity and Integration Protocol: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MCIP

Trigger	Name the condition that activates MCIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-25 - Moussaif Complementarity and Integration Protocol.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-25 - Moussaif Complementarity and Integration Protocol: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MCIP is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-25 - Moussaif Complementarity and Integration Protocol: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: complementarity matrix; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-25 - Moussaif Complementarity and Integration Protocol: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-25 - Moussaif Complementarity and Integration Protocol: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MCIP: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-25 - Moussaif Complementarity and Integration Protocol: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure marginal capability gain, redundancy, interface defects. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-25 - Moussaif Complementarity and Integration Protocol: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by components combined because they are individually impressive rather than jointly useful, the baseline produces an answer directly. The MCIP route emits complementarity matrix, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-25 - Moussaif Complementarity and Integration Protocol: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MCIP

Trigger	Name the condition that activates MCIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-25 - Moussaif Complementarity and Integration Protocol.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-25 - Moussaif Complementarity and Integration Protocol: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where integration is justified only when components add non-redundant causal value. Match models, prompts, budget and tools. If MCIP does not improve marginal capability gain, redundancy, interface defects, the claimed mechanism is not supported under that regime. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-25 - Moussaif Complementarity and Integration Protocol: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: false synergy and interface complexity. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-25 - Moussaif Complementarity and Integration Protocol: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MCIP: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MASO, MCD, MSSDT), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-25 - Moussaif Complementarity and Integration Protocol: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: marginal capability gain, redundancy, interface defects. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-25 - Moussaif Complementarity and Integration Protocol: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCIP

Trigger	Name the condition that activates MCIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-25 - Moussaif Complementarity and Integration Protocol.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-26 - Moussaif Court-Independence Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: verification captured by the assumptions and incentives of generation. Core claim: courts need independent prompts, evidence access and verdict authority. Operational claim analysis for TH-26 - Moussaif Court-Independence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research thesis, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

PRIMARY OPERATIONAL OBJECT

The required artifact is court independence report. A verbal explanation without this object is not activation of MCIT. The object must be schema-valid, versioned and linked to the task contract. Boundary analysis for TH-26 - Moussaif Court-Independence Theory: define inclusion, exclusion and abstention regions for primary operational object. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DEPENDENCIES

Principal dependencies and neighboring constructs: MDFC, MGFC, MDBCT. Integration should preserve separate responsibilities and permit component-level ablation. Artifact discipline for TH-26 - Moussaif Court-Independence Theory: the dependencies claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

RESEARCH QUESTIONS

Can MCIT improve error diversity, correlated-failure rate, override audit under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Integration analysis for TH-26 - Moussaif Court-Independence Theory: map research questions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MCIT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets verification captured by the assumptions and incentives of generation. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Boundary analysis for TH-26 - Moussaif Court-Independence Theory: define inclusion, exclusion and abstention regions for governing problem. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BOUNDARY AND EXCLUSIONS

MCIT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Artifact discipline for TH-26 - Moussaif Court-Independence Theory: the boundary and exclusions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

KNOWN RISK

Primary conceptual risk: nominal independence with shared blind spots. The registry must record counterexamples, unresolved definitions and alternative explanations. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate known risk into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MCIT

Trigger	Name the condition that activates MCIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-26 - Moussaif Court-Independence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-26 - Moussaif Court-Independence Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MCIT transforms an admissible state into court independence report plus an uncertainty and provenance record. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate state and notation into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Boundary analysis for TH-26 - Moussaif Court-Independence Theory: define inclusion, exclusion and abstention regions for invariants. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Risk and counterexample analysis for TH-26 - Moussaif Court-Independence Theory: turn the named limitation in complexity model into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid court independence report. Integration analysis for TH-26 - Moussaif Court-Independence Theory: map formal failure to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MCIT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit court independence report; 7) write telemetry and residual uncertainty. Risk and counterexample analysis for TH-26 - Moussaif Court-Independence Theory: turn the named limitation in mechanism into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DECISION RULE

Adopt activation when expected gain in error diversity, correlated-failure rate, override audit exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Boundary analysis for TH-26 - Moussaif Court-Independence Theory: define inclusion, exclusion and abstention regions for decision rule. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MDFC, MGFC, MDBCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Integration analysis for TH-26 - Moussaif Court-Independence Theory: map interfaces to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MCIT

Trigger	Name the condition that activates MCIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-26 - Moussaif Court-Independence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-26 - Moussaif Court-Independence Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MCIT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Operational claim analysis for TH-26 - Moussaif Court-Independence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For runtime architecture, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OUTPUT CONTRACT

Output: court independence report; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Operational claim analysis for TH-26 - Moussaif Court-Independence Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For output contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Risk and counterexample analysis for TH-26 - Moussaif Court-Independence Theory: turn the named limitation in security and authority into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate integration verdicts into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MCIT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Integration analysis for TH-26 - Moussaif Court-Independence Theory: map input contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure error diversity, correlated-failure rate, override audit. Sensitive content is redacted or referenced, not copied into telemetry. Research-program analysis for TH-26 - Moussaif Court-Independence Theory: decompose observability into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

WORKED SCENARIO

Given a task dominated by verification captured by the assumptions and incentives of generation, the baseline produces an answer directly. The MCIT route emits court independence report, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Artifact discipline for TH-26 - Moussaif Court-Independence Theory: the worked scenario claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MCIT

Trigger	Name the condition that activates MCIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-26 - Moussaif Court-Independence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-26 - Moussaif Court-Independence Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where courts need independent prompts, evidence access and verdict authority. Match models, prompts, budget and tools. If MCIT does not improve error diversity, correlated-failure rate, override audit, the claimed mechanism is not supported under that regime. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate primary falsifier into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Boundary analysis for TH-26 - Moussaif Court-Independence Theory: define inclusion, exclusion and abstention regions for baselines. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: nominal independence with shared blind spots. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate adversarial cases into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Integration analysis for TH-26 - Moussaif Court-Independence Theory: map next research action to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MCIT: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MDFC, MGFC, MDBCT), the court and the memory write separately. Measure both capability loss and cost reduction. Integration analysis for TH-26 - Moussaif Court-Independence Theory: map ablation plan to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

METRICS

Primary metrics: error diversity, correlated-failure rate, override audit. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate metrics into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Mechanism specification for TH-26 - Moussaif Court-Independence Theory: translate evidence ladder into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MCIT

Trigger	Name the condition that activates MCIT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-26 - Moussaif Court-Independence Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-27 - Moussaif Model-Portfolio Router Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: one model used across tasks with different reasoning, privacy and latency needs. Core claim: a capability-profiled portfolio should route by task, risk, cost and evidence needs. Mechanism specification for TH-27 - Moussaif Model-Portfolio Router Theory: translate research thesis into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

PRIMARY OPERATIONAL OBJECT

The required artifact is model selection packet. A verbal explanation without this object is not activation of MMPR. The object must be schema-valid, versioned and linked to the task contract. Integration analysis for TH-27 - Moussaif Model-Portfolio Router Theory: map primary operational object to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DEPENDENCIES

Principal dependencies and neighboring constructs: MSHR, MCVT, MTEL. Integration should preserve separate responsibilities and permit component-level ablation. Boundary analysis for TH-27 - Moussaif Model-Portfolio Router Theory: define inclusion, exclusion and abstention regions for dependencies. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

RESEARCH QUESTIONS

Can MMPR improve quality, cost, privacy, fallback success under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Risk and counterexample analysis for TH-27 - Moussaif Model-Portfolio Router Theory: turn the named limitation in research questions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MMPR: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets one model used across tasks with different reasoning, privacy and latency needs. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Integration analysis for TH-27 - Moussaif Model-Portfolio Router Theory: map governing problem to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BOUNDARY AND EXCLUSIONS

MMPR does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Boundary analysis for TH-27 - Moussaif Model-Portfolio Router Theory: define inclusion, exclusion and abstention regions for boundary and exclusions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

KNOWN RISK

Primary conceptual risk: benchmark gaming and provider lock-in. The registry must record counterexamples, unresolved definitions and alternative explanations. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the known risk claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MMPR

Trigger	Name the condition that activates MMPR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-27 - Moussaif Model-Portfolio Router Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-27 - Moussaif Model-Portfolio Router Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MMPR transforms an admissible state into model selection packet plus an uncertainty and provenance record. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the state and notation claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Integration analysis for TH-27 - Moussaif Model-Portfolio Router Theory: map invariants to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Research-program analysis for TH-27 - Moussaif Model-Portfolio Router Theory: decompose complexity model into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid model selection packet. Risk and counterexample analysis for TH-27 - Moussaif Model-Portfolio Router Theory: turn the named limitation in formal failure into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MMPR: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit model selection packet; 7) write telemetry and residual uncertainty. Research-program analysis for TH-27 - Moussaif Model-Portfolio Router Theory: decompose mechanism into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DECISION RULE

Adopt activation when expected gain in quality, cost, privacy, fallback success exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Integration analysis for TH-27 - Moussaif Model-Portfolio Router Theory: map decision rule to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MSHR, MCVT, MTEL. Outputs feed planning, routing, courts, deployment or memory according to the object type. Risk and counterexample analysis for TH-27 - Moussaif Model-Portfolio Router Theory: turn the named limitation in interfaces into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MMPR

Trigger	Name the condition that activates MMPR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-27 - Moussaif Model-Portfolio Router Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-27 - Moussaif Model-Portfolio Router Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MMPR is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Mechanism specification for TH-27 - Moussaif Model-Portfolio Router Theory: translate runtime architecture into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OUTPUT CONTRACT

Output: model selection packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Mechanism specification for TH-27 - Moussaif Model-Portfolio Router Theory: translate output contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Research-program analysis for TH-27 - Moussaif Model-Portfolio Router Theory: decompose security and authority into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the integration verdicts claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MMPR: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Risk and counterexample analysis for TH-27 - Moussaif Model-Portfolio Router Theory: turn the named limitation in input contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure quality, cost, privacy, fallback success. Sensitive content is redacted or referenced, not copied into telemetry. Operational claim analysis for TH-27 - Moussaif Model-Portfolio Router Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For observability, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

WORKED SCENARIO

Given a task dominated by one model used across tasks with different reasoning, privacy and latency needs, the baseline produces an answer directly. The MMPR route emits model selection packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Boundary analysis for TH-27 - Moussaif Model-Portfolio Router Theory: define inclusion, exclusion and abstention regions for worked scenario. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MMPR

Trigger	Name the condition that activates MMPR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-27 - Moussaif Model-Portfolio Router Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-27 - Moussaif Model-Portfolio Router Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where a capability-profiled portfolio should route by task, risk, cost and evidence needs. Match models, prompts, budget and tools. If MMPR does not improve quality, cost, privacy, fallback success, the claimed mechanism is not supported under that regime. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the primary falsifier claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Integration analysis for TH-27 - Moussaif Model-Portfolio Router Theory: map baselines to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: benchmark gaming and provider lock-in. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the adversarial cases claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Risk and counterexample analysis for TH-27 - Moussaif Model-Portfolio Router Theory: turn the named limitation in next research action into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MMPR: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MSHR, MCVT, MTEL), the court and the memory write separately. Measure both capability loss and cost reduction. Risk and counterexample analysis for TH-27 - Moussaif Model-Portfolio Router Theory: turn the named limitation in ablation plan into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

METRICS

Primary metrics: quality, cost, privacy, fallback success. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the metrics claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Artifact discipline for TH-27 - Moussaif Model-Portfolio Router Theory: the evidence ladder claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MMPR

Trigger	Name the condition that activates MMPR; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-27 - Moussaif Model-Portfolio Router Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-28 - Moussaif Token-Economy Law: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: token growth mistaken for intelligence gain. Core claim: tokens should be allocated where marginal evidence or decision value is positive. Research-program analysis for TH-28 - Moussaif Token-Economy Law: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is token budget ledger. A verbal explanation without this object is not activation of MTEL. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-28 - Moussaif Token-Economy Law: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: Q-MCD, MCVT, MMPR. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-28 - Moussaif Token-Economy Law: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MTEL improve value per token, compression loss, budget overrun under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-28 - Moussaif Token-Economy Law: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MTEL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets token growth mistaken for intelligence gain. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-28 - Moussaif Token-Economy Law: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MTEL does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-28 - Moussaif Token-Economy Law: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: over-optimization that suppresses useful exploration. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MTEL

Trigger	Name the condition that activates MTEL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-28 - Moussaif Token-Economy Law.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-28 - Moussaif Token-Economy Law: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MTEL transforms an admissible state into token budget ledger plus an uncertainty and provenance record. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-28 - Moussaif Token-Economy Law: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-28 - Moussaif Token-Economy Law: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid token budget ledger. Boundary analysis for TH-28 - Moussaif Token-Economy Law: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MTEL: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit token budget ledger; 7) write telemetry and residual uncertainty. Integration analysis for TH-28 - Moussaif Token-Economy Law: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in value per token, compression loss, budget overrun exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-28 - Moussaif Token-Economy Law: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and Q-MCD, MCVT, MMPR. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-28 - Moussaif Token-Economy Law: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MTEL

Trigger	Name the condition that activates MTEL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-28 - Moussaif Token-Economy Law.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-28 - Moussaif Token-Economy Law: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MTEL is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-28 - Moussaif Token-Economy Law: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: token budget ledger; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-28 - Moussaif Token-Economy Law: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-28 - Moussaif Token-Economy Law: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MTEL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-28 - Moussaif Token-Economy Law: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure value per token, compression loss, budget overrun. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-28 - Moussaif Token-Economy Law: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by token growth mistaken for intelligence gain, the baseline produces an answer directly. The MTEL route emits token budget ledger, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-28 - Moussaif Token-Economy Law: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MTEL

Trigger	Name the condition that activates MTEL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-28 - Moussaif Token-Economy Law.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-28 - Moussaif Token-Economy Law: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where tokens should be allocated where marginal evidence or decision value is positive. Match models, prompts, budget and tools. If MTEL does not improve value per token, compression loss, budget overrun, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-28 - Moussaif Token-Economy Law: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: over-optimization that suppresses useful exploration. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-28 - Moussaif Token-Economy Law: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MTEL: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (Q-MCD, MCVT, MMPR), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-28 - Moussaif Token-Economy Law: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: value per token, compression loss, budget overrun. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-28 - Moussaif Token-Economy Law: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MTEL

Trigger	Name the condition that activates MTEL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-28 - Moussaif Token-Economy Law.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-29 - Moussaif Boundary-Resilience and RPM Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: systems failing under rate, context, permission and interface boundaries. Core claim: resilience emerges from explicit boundary models, backpressure and graceful degradation. Risk and counterexample analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: turn the named limitation in research thesis into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

PRIMARY OPERATIONAL OBJECT

The required artifact is boundary and RPM control plan. A verbal explanation without this object is not activation of MBRRT. The object must be schema-valid, versioned and linked to the task contract. Mechanism specification for TH-29 - Moussaif Boundary-Resilience and RPM Theory: translate primary operational object into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DEPENDENCIES

Principal dependencies and neighboring constructs: MTEL, MAECT, MRRMT. Integration should preserve separate responsibilities and permit component-level ablation. Operational claim analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For dependencies, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

RESEARCH QUESTIONS

Can MBRRT improve throughput, degradation quality, retry amplification under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Artifact discipline for TH-29 - Moussaif Boundary-Resilience and RPM Theory: the research questions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MBRRT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets systems failing under rate, context, permission and interface boundaries. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Mechanism specification for TH-29 - Moussaif Boundary-Resilience and RPM Theory: translate governing problem into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BOUNDARY AND EXCLUSIONS

MBRRT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Operational claim analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For boundary and exclusions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

KNOWN RISK

Primary conceptual risk: retry storms and hidden quota coupling. The registry must record counterexamples, unresolved definitions and alternative explanations. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose known risk into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MBRRT

Trigger	Name the condition that activates MBRRT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-29 - Moussaif Boundary-Resilience and RPM Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-29 - Moussaif Boundary-Resilience and RPM Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MBRRT transforms an admissible state into boundary and RPM control plan plus an uncertainty and provenance record. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose state and notation into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Mechanism specification for TH-29 - Moussaif Boundary-Resilience and RPM Theory: translate invariants into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Boundary analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: define inclusion, exclusion and abstention regions for complexity model. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid boundary and RPM control plan. Artifact discipline for TH-29 - Moussaif Boundary-Resilience and RPM Theory: the formal failure claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MBRRT: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit boundary and RPM control plan; 7) write telemetry and residual uncertainty. Boundary analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: define inclusion, exclusion and abstention regions for mechanism. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DECISION RULE

Adopt activation when expected gain in throughput, degradation quality, retry amplification exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Mechanism specification for TH-29 - Moussaif Boundary-Resilience and RPM Theory: translate decision rule into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MTEL, MAECT, MRRMT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Artifact discipline for TH-29 - Moussaif Boundary-Resilience and RPM Theory: the interfaces claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MBRRT

Trigger	Name the condition that activates MBRRT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-29 - Moussaif Boundary-Resilience and RPM Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-29 - Moussaif Boundary-Resilience and RPM Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MBRRT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Risk and counterexample analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: turn the named limitation in runtime architecture into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OUTPUT CONTRACT

Output: boundary and RPM control plan; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Risk and counterexample analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: turn the named limitation in output contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Boundary analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: define inclusion, exclusion and abstention regions for security and authority. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose integration verdicts into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MBRRT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Artifact discipline for TH-29 - Moussaif Boundary-Resilience and RPM Theory: the input contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure throughput, degradation quality, retry amplification. Sensitive content is redacted or referenced, not copied into telemetry. Integration analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: map observability to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

WORKED SCENARIO

Given a task dominated by systems failing under rate, context, permission and interface boundaries, the baseline produces an answer directly. The MBRRT route emits boundary and RPM control plan, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Operational claim analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For worked scenario, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MBRRT

Trigger	Name the condition that activates MBRRT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-29 - Moussaif Boundary-Resilience and RPM Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-29 - Moussaif Boundary-Resilience and RPM Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where resilience emerges from explicit boundary models, backpressure and graceful degradation. Match models, prompts, budget and tools. If MBRRT does not improve throughput, degradation quality, retry amplification, the claimed mechanism is not supported under that regime. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose primary falsifier into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Mechanism specification for TH-29 - Moussaif Boundary-Resilience and RPM Theory: translate baselines into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: retry storms and hidden quota coupling. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose adversarial cases into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Artifact discipline for TH-29 - Moussaif Boundary-Resilience and RPM Theory: the next research action claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MBRRT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MTEL, MAECT, MRRMT), the court and the memory write separately. Measure both capability loss and cost reduction. Artifact discipline for TH-29 - Moussaif Boundary-Resilience and RPM Theory: the ablation plan claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

METRICS

Primary metrics: throughput, degradation quality, retry amplification. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose metrics into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Research-program analysis for TH-29 - Moussaif Boundary-Resilience and RPM Theory: decompose evidence ladder into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MBRRT

Trigger	Name the condition that activates MBRRT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-29 - Moussaif Boundary-Resilience and RPM Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-30 - Moussaif Epistemic-Calibration Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: confidence expressed without relation to empirical correctness. Core claim: claims should carry calibrated uncertainty and evidence-grade-specific language. Boundary analysis for TH-30 - Moussaif Epistemic-Calibration Theory: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is calibrated claim packet. A verbal explanation without this object is not activation of MECT. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-30 - Moussaif Epistemic-Calibration Theory: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: MELT, MDFC, MDBCT. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-30 - Moussaif Epistemic-Calibration Theory: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MECT improve Brier score, reliability error, abstention quality under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-30 - Moussaif Epistemic-Calibration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MECT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets confidence expressed without relation to empirical correctness. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-30 - Moussaif Epistemic-Calibration Theory: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MECT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-30 - Moussaif Epistemic-Calibration Theory: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: cosmetic hedging and miscalibrated rare events. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MECT

Trigger	Name the condition that activates MECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-30 - Moussaif Epistemic-Calibration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-30 - Moussaif Epistemic-Calibration Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MECT transforms an admissible state into calibrated claim packet plus an uncertainty and provenance record. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-30 - Moussaif Epistemic-Calibration Theory: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-30 - Moussaif Epistemic-Calibration Theory: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid calibrated claim packet. Operational claim analysis for TH-30 - Moussaif Epistemic-Calibration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MECT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit calibrated claim packet; 7) write telemetry and residual uncertainty. Mechanism specification for TH-30 - Moussaif Epistemic-Calibration Theory: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in Brier score, reliability error, abstention quality exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-30 - Moussaif Epistemic-Calibration Theory: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MELT, MDFC, MDBCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-30 - Moussaif Epistemic-Calibration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MECT

Trigger	Name the condition that activates MECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-30 - Moussaif Epistemic-Calibration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-30 - Moussaif Epistemic-Calibration Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MECT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-30 - Moussaif Epistemic-Calibration Theory: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: calibrated claim packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-30 - Moussaif Epistemic-Calibration Theory: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-30 - Moussaif Epistemic-Calibration Theory: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MECT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-30 - Moussaif Epistemic-Calibration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure Brier score, reliability error, abstention quality. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-30 - Moussaif Epistemic-Calibration Theory: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by confidence expressed without relation to empirical correctness, the baseline produces an answer directly. The MECT route emits calibrated claim packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-30 - Moussaif Epistemic-Calibration Theory: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MECT

Trigger	Name the condition that activates MECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-30 - Moussaif Epistemic-Calibration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-30 - Moussaif Epistemic-Calibration Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where claims should carry calibrated uncertainty and evidence-grade-specific language. Match models, prompts, budget and tools. If MECT does not improve Brier score, reliability error, abstention quality, the claimed mechanism is not supported under that regime. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-30 - Moussaif Epistemic-Calibration Theory: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: cosmetic hedging and miscalibrated rare events. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-30 - Moussaif Epistemic-Calibration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MECT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MELT, MDFC, MDBCT), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-30 - Moussaif Epistemic-Calibration Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: Brier score, reliability error, abstention quality. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-30 - Moussaif Epistemic-Calibration Theory: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MECT

Trigger	Name the condition that activates MECT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-30 - Moussaif Epistemic-Calibration Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-31 - Moussaif Authority-Boundary Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: autonomous execution beyond granted permission. Core claim: authority must be explicit, scoped, revocable and checked at action time. Artifact discipline for TH-31 - Moussaif Authority-Boundary Theory: the research thesis claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

PRIMARY OPERATIONAL OBJECT

The required artifact is authority token and decision log. A verbal explanation without this object is not activation of MABT. The object must be schema-valid, versioned and linked to the task contract. Risk and counterexample analysis for TH-31 - Moussaif Authority-Boundary Theory: turn the named limitation in primary operational object into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DEPENDENCIES

Principal dependencies and neighboring constructs: MCIL, MTIT, MIRTT. Integration should preserve separate responsibilities and permit component-level ablation. Integration analysis for TH-31 - Moussaif Authority-Boundary Theory: map dependencies to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

RESEARCH QUESTIONS

Can MABT improve unauthorized action rate, revocation latency, auditability under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Research-program analysis for TH-31 - Moussaif Authority-Boundary Theory: decompose research questions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MABT: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets autonomous execution beyond granted permission. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Risk and counterexample analysis for TH-31 - Moussaif Authority-Boundary Theory: turn the named limitation in governing problem into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BOUNDARY AND EXCLUSIONS

MABT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Integration analysis for TH-31 - Moussaif Authority-Boundary Theory: map boundary and exclusions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

KNOWN RISK

Primary conceptual risk: permission ambiguity and unsafe implied consent. The registry must record counterexamples, unresolved definitions and alternative explanations. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for known risk. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MABT

Trigger	Name the condition that activates MABT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-31 - Moussaif Authority-Boundary Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-31 - Moussaif Authority-Boundary Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MABT transforms an admissible state into authority token and decision log plus an uncertainty and provenance record. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for state and notation. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Risk and counterexample analysis for TH-31 - Moussaif Authority-Boundary Theory: turn the named limitation in invariants into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Operational claim analysis for TH-31 - Moussaif Authority-Boundary Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For complexity model, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid authority token and decision log. Research-program analysis for TH-31 - Moussaif Authority-Boundary Theory: decompose formal failure into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MABT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit authority token and decision log; 7) write telemetry and residual uncertainty. Operational claim analysis for TH-31 - Moussaif Authority-Boundary Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For mechanism, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DECISION RULE

Adopt activation when expected gain in unauthorized action rate, revocation latency, auditability exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Risk and counterexample analysis for TH-31 - Moussaif Authority-Boundary Theory: turn the named limitation in decision rule into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MCIL, MTIT, MIRTT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Research-program analysis for TH-31 - Moussaif Authority-Boundary Theory: decompose interfaces into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MABT

Trigger	Name the condition that activates MABT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-31 - Moussaif Authority-Boundary Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-31 - Moussaif Authority-Boundary Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MABT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Artifact discipline for TH-31 - Moussaif Authority-Boundary Theory: the runtime architecture claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OUTPUT CONTRACT

Output: authority token and decision log; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Artifact discipline for TH-31 - Moussaif Authority-Boundary Theory: the output contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Operational claim analysis for TH-31 - Moussaif Authority-Boundary Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For security and authority, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for integration verdicts. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MABT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Research-program analysis for TH-31 - Moussaif Authority-Boundary Theory: decompose input contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure unauthorized action rate, revocation latency, auditability. Sensitive content is redacted or referenced, not copied into telemetry. Mechanism specification for TH-31 - Moussaif Authority-Boundary Theory: translate observability into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

WORKED SCENARIO

Given a task dominated by autonomous execution beyond granted permission, the baseline produces an answer directly. The MABT route emits authority token and decision log, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Integration analysis for TH-31 - Moussaif Authority-Boundary Theory: map worked scenario to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MABT

Trigger	Name the condition that activates MABT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-31 - Moussaif Authority-Boundary Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-31 - Moussaif Authority-Boundary Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where authority must be explicit, scoped, revocable and checked at action time. Match models, prompts, budget and tools. If MABT does not improve unauthorized action rate, revocation latency, auditability, the claimed mechanism is not supported under that regime. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for primary falsifier. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Risk and counterexample analysis for TH-31 - Moussaif Authority-Boundary Theory: turn the named limitation in baselines into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: permission ambiguity and unsafe implied consent. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for adversarial cases. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Research-program analysis for TH-31 - Moussaif Authority-Boundary Theory: decompose next research action into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MABT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MCIL, MTIT, MIRTT), the court and the memory write separately. Measure both capability loss and cost reduction. Research-program analysis for TH-31 - Moussaif Authority-Boundary Theory: decompose ablation plan into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

METRICS

Primary metrics: unauthorized action rate, revocation latency, auditability. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for metrics. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Boundary analysis for TH-31 - Moussaif Authority-Boundary Theory: define inclusion, exclusion and abstention regions for evidence ladder. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MABT

Trigger	Name the condition that activates MABT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-31 - Moussaif Authority-Boundary Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-32 - Moussaif Injection-Resistant Toolchain Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: untrusted content steering privileged tool execution. Core claim: instruction and data channels must be isolated with taint tracking and least privilege. Integration analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: map research thesis to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

PRIMARY OPERATIONAL OBJECT

The required artifact is taint-aware tool plan. A verbal explanation without this object is not activation of MIRTT. The object must be schema-valid, versioned and linked to the task contract. Operational claim analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary operational object, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DEPENDENCIES

Principal dependencies and neighboring constructs: MABT, MTIT, MDST. Integration should preserve separate responsibilities and permit component-level ablation. Research-program analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: decompose dependencies into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

RESEARCH QUESTIONS

Can MIRTT improve attack success, detection, containment, recovery under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Mechanism specification for TH-32 - Moussaif Injection-Resistant Toolchain Theory: translate research questions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MIRTT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets untrusted content steering privileged tool execution. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Operational claim analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For governing problem, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BOUNDARY AND EXCLUSIONS

MIRTT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Research-program analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: decompose boundary and exclusions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

KNOWN RISK

Primary conceptual risk: novel injection paths and overblocking benign content. The registry must record counterexamples, unresolved definitions and alternative explanations. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in known risk into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MIRTT

Trigger	Name the condition that activates MIRTT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-32 - Moussaif Injection-Resistant Toolchain Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-32 - Moussaif Injection-Resistant Toolchain Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MIRTT transforms an admissible state into taint-aware tool plan plus an uncertainty and provenance record. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in state and notation into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Operational claim analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For invariants, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Artifact discipline for TH-32 - Moussaif Injection-Resistant Toolchain Theory: the complexity model claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid taint-aware tool plan. Mechanism specification for TH-32 - Moussaif Injection-Resistant Toolchain Theory: translate formal failure into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MIRTT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit taint-aware tool plan; 7) write telemetry and residual uncertainty. Artifact discipline for TH-32 - Moussaif Injection-Resistant Toolchain Theory: the mechanism claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DECISION RULE

Adopt activation when expected gain in attack success, detection, containment, recovery exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Operational claim analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For decision rule, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MABT, MTIT, MDST. Outputs feed planning, routing, courts, deployment or memory according to the object type. Mechanism specification for TH-32 - Moussaif Injection-Resistant Toolchain Theory: translate interfaces into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MIRTT

Trigger	Name the condition that activates MIRTT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-32 - Moussaif Injection-Resistant Toolchain Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-32 - Moussaif Injection-Resistant Toolchain Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MIRTT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Integration analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: map runtime architecture to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OUTPUT CONTRACT

Output: taint-aware tool plan; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Integration analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: map output contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Artifact discipline for TH-32 - Moussaif Injection-Resistant Toolchain Theory: the security and authority claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in integration verdicts into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MIRTT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Mechanism specification for TH-32 - Moussaif Injection-Resistant Toolchain Theory: translate input contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure attack success, detection, containment, recovery. Sensitive content is redacted or referenced, not copied into telemetry. Boundary analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: define inclusion, exclusion and abstention regions for observability. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

WORKED SCENARIO

Given a task dominated by untrusted content steering privileged tool execution, the baseline produces an answer directly. The MIRTT route emits taint-aware tool plan, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Research-program analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: decompose worked scenario into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MIRTT

Trigger	Name the condition that activates MIRTT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-32 - Moussaif Injection-Resistant Toolchain Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-32 - Moussaif Injection-Resistant Toolchain Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where instruction and data channels must be isolated with taint tracking and least privilege. Match models, prompts, budget and tools. If MIRTT does not improve attack success, detection, containment, recovery, the claimed mechanism is not supported under that regime. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in primary falsifier into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Operational claim analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For baselines, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: novel injection paths and overblocking benign content. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in adversarial cases into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Mechanism specification for TH-32 - Moussaif Injection-Resistant Toolchain Theory: translate next research action into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MIRTT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MABT, MTIT, MDST), the court and the memory write separately. Measure both capability loss and cost reduction. Mechanism specification for TH-32 - Moussaif Injection-Resistant Toolchain Theory: translate ablation plan into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

METRICS

Primary metrics: attack success, detection, containment, recovery. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in metrics into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Risk and counterexample analysis for TH-32 - Moussaif Injection-Resistant Toolchain Theory: turn the named limitation in evidence ladder into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MIRTT

Trigger	Name the condition that activates MIRTT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-32 - Moussaif Injection-Resistant Toolchain Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-33 - Moussaif Data-Sovereignty Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: data leaving intended legal, privacy or operational boundaries. Core claim: data use should be purpose-bound, minimized, localized and traceable. Risk and counterexample analysis for TH-33 - Moussaif Data-Sovereignty Theory: turn the named limitation in research thesis into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

PRIMARY OPERATIONAL OBJECT

The required artifact is data sovereignty manifest. A verbal explanation without this object is not activation of MDST. The object must be schema-valid, versioned and linked to the task contract. Mechanism specification for TH-33 - Moussaif Data-Sovereignty Theory: translate primary operational object into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DEPENDENCIES

Principal dependencies and neighboring constructs: MABT, MIRTT, MAOLT. Integration should preserve separate responsibilities and permit component-level ablation. Operational claim analysis for TH-33 - Moussaif Data-Sovereignty Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For dependencies, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

RESEARCH QUESTIONS

Can MDST improve data exposure, residency compliance, deletion proof under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Artifact discipline for TH-33 - Moussaif Data-Sovereignty Theory: the research questions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MDST: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets data leaving intended legal, privacy or operational boundaries. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Mechanism specification for TH-33 - Moussaif Data-Sovereignty Theory: translate governing problem into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BOUNDARY AND EXCLUSIONS

MDST does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Operational claim analysis for TH-33 - Moussaif Data-Sovereignty Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For boundary and exclusions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

KNOWN RISK

Primary conceptual risk: policy complexity and incomplete lineage. The registry must record counterexamples, unresolved definitions and alternative explanations. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose known risk into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MDST

Trigger	Name the condition that activates MDST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-33 - Moussaif Data-Sovereignty Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-33 - Moussaif Data-Sovereignty Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MDST transforms an admissible state into data sovereignty manifest plus an uncertainty and provenance record. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose state and notation into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Mechanism specification for TH-33 - Moussaif Data-Sovereignty Theory: translate invariants into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Boundary analysis for TH-33 - Moussaif Data-Sovereignty Theory: define inclusion, exclusion and abstention regions for complexity model. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid data sovereignty manifest. Artifact discipline for TH-33 - Moussaif Data-Sovereignty Theory: the formal failure claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MDST: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit data sovereignty manifest; 7) write telemetry and residual uncertainty. Boundary analysis for TH-33 - Moussaif Data-Sovereignty Theory: define inclusion, exclusion and abstention regions for mechanism. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DECISION RULE

Adopt activation when expected gain in data exposure, residency compliance, deletion proof exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Mechanism specification for TH-33 - Moussaif Data-Sovereignty Theory: translate decision rule into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MABT, MIRTT, MAOLT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Artifact discipline for TH-33 - Moussaif Data-Sovereignty Theory: the interfaces claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MDST

Trigger	Name the condition that activates MDST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-33 - Moussaif Data-Sovereignty Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-33 - Moussaif Data-Sovereignty Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MDST is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Risk and counterexample analysis for TH-33 - Moussaif Data-Sovereignty Theory: turn the named limitation in runtime architecture into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OUTPUT CONTRACT

Output: data sovereignty manifest; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Risk and counterexample analysis for TH-33 - Moussaif Data-Sovereignty Theory: turn the named limitation in output contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Boundary analysis for TH-33 - Moussaif Data-Sovereignty Theory: define inclusion, exclusion and abstention regions for security and authority. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose integration verdicts into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MDST: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Artifact discipline for TH-33 - Moussaif Data-Sovereignty Theory: the input contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure data exposure, residency compliance, deletion proof. Sensitive content is redacted or referenced, not copied into telemetry. Integration analysis for TH-33 - Moussaif Data-Sovereignty Theory: map observability to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

WORKED SCENARIO

Given a task dominated by data leaving intended legal, privacy or operational boundaries, the baseline produces an answer directly. The MDST route emits data sovereignty manifest, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Operational claim analysis for TH-33 - Moussaif Data-Sovereignty Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For worked scenario, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MDST

Trigger	Name the condition that activates MDST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-33 - Moussaif Data-Sovereignty Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-33 - Moussaif Data-Sovereignty Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where data use should be purpose-bound, minimized, localized and traceable. Match models, prompts, budget and tools. If MDST does not improve data exposure, residency compliance, deletion proof, the claimed mechanism is not supported under that regime. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose primary falsifier into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Mechanism specification for TH-33 - Moussaif Data-Sovereignty Theory: translate baselines into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: policy complexity and incomplete lineage. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose adversarial cases into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Artifact discipline for TH-33 - Moussaif Data-Sovereignty Theory: the next research action claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MDST: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MABT, MIRTT, MAOLT), the court and the memory write separately. Measure both capability loss and cost reduction. Artifact discipline for TH-33 - Moussaif Data-Sovereignty Theory: the ablation plan claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

METRICS

Primary metrics: data exposure, residency compliance, deletion proof. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose metrics into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Research-program analysis for TH-33 - Moussaif Data-Sovereignty Theory: decompose evidence ladder into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MDST

Trigger	Name the condition that activates MDST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-33 - Moussaif Data-Sovereignty Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-34 - Moussaif Agentic Observability-Ledger Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: agent decisions and world effects that cannot be reconstructed. Core claim: critical state transitions require a tamper-evident, privacy-aware ledger. Risk and counterexample analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: turn the named limitation in research thesis into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

PRIMARY OPERATIONAL OBJECT

The required artifact is observability ledger. A verbal explanation without this object is not activation of MAOLT. The object must be schema-valid, versioned and linked to the task contract. Mechanism specification for TH-34 - Moussaif Agentic Observability-Ledger Theory: translate primary operational object into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DEPENDENCIES

Principal dependencies and neighboring constructs: MMT, MWSM, MRRMT. Integration should preserve separate responsibilities and permit component-level ablation. Operational claim analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For dependencies, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

RESEARCH QUESTIONS

Can MAOLT improve trace coverage, diagnostic time, privacy leakage under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Artifact discipline for TH-34 - Moussaif Agentic Observability-Ledger Theory: the research questions claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MAOLT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets agent decisions and world effects that cannot be reconstructed. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Mechanism specification for TH-34 - Moussaif Agentic Observability-Ledger Theory: translate governing problem into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

BOUNDARY AND EXCLUSIONS

MAOLT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Operational claim analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For boundary and exclusions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

KNOWN RISK

Primary conceptual risk: telemetry overload and sensitive trace retention. The registry must record counterexamples, unresolved definitions and alternative explanations. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose known risk into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MAOLT

Trigger	Name the condition that activates MAOLT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-34 - Moussaif Agentic Observability-Ledger Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-34 - Moussaif Agentic Observability-Ledger Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MAOLT transforms an admissible state into observability ledger plus an uncertainty and provenance record. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose state and notation into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Mechanism specification for TH-34 - Moussaif Agentic Observability-Ledger Theory: translate invariants into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Boundary analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: define inclusion, exclusion and abstention regions for complexity model. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid observability ledger. Artifact discipline for TH-34 - Moussaif Agentic Observability-Ledger Theory: the formal failure claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MAOLT: budgetو baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit observability ledger; 7) write telemetry and residual uncertainty. Boundary analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: define inclusion, exclusion and abstention regions for mechanism. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

DECISION RULE

Adopt activation when expected gain in trace coverage, diagnostic time, privacy leakage exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Mechanism specification for TH-34 - Moussaif Agentic Observability-Ledger Theory: translate decision rule into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MMT, MWSM, MRRMT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Artifact discipline for TH-34 - Moussaif Agentic Observability-Ledger Theory: the interfaces claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MAOLT

Trigger	Name the condition that activates MAOLT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-34 - Moussaif Agentic Observability-Ledger Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-34 - Moussaif Agentic Observability-Ledger Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MAOLT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Risk and counterexample analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: turn the named limitation in runtime architecture into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OUTPUT CONTRACT

Output: observability ledger; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Risk and counterexample analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: turn the named limitation in output contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Boundary analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: define inclusion, exclusion and abstention regions for security and authority. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose integration verdicts into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MAOLT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Artifact discipline for TH-34 - Moussaif Agentic Observability-Ledger Theory: the input contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure trace coverage, diagnostic time, privacy leakage. Sensitive content is redacted or referenced, not copied into telemetry. Integration analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: map observability to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

WORKED SCENARIO

Given a task dominated by agent decisions and world effects that cannot be reconstructed, the baseline produces an answer directly. The MAOLT route emits observability ledger, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Operational claim analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For worked scenario, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MAOLT

Trigger	Name the condition that activates MAOLT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-34 - Moussaif Agentic Observability-Ledger Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-34 - Moussaif Agentic Observability-Ledger Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where critical state transitions require a tamper-evident, privacy-aware ledger. Match models, prompts, budget and tools. If MAOLT does not improve trace coverage, diagnostic time, privacy leakage, the claimed mechanism is not supported under that regime. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose primary falsifier into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Mechanism specification for TH-34 - Moussaif Agentic Observability-Ledger Theory: translate baselines into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: telemetry overload and sensitive trace retention. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose adversarial cases into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Artifact discipline for TH-34 - Moussaif Agentic Observability-Ledger Theory: the next research action claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MAOLT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MMT, MWSM, MRRMT), the court and the memory write separately. Measure both capability loss and cost reduction. Artifact discipline for TH-34 - Moussaif Agentic Observability-Ledger Theory: the ablation plan claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

METRICS

Primary metrics: trace coverage, diagnostic time, privacy leakage. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose metrics into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Research-program analysis for TH-34 - Moussaif Agentic Observability-Ledger Theory: decompose evidence ladder into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MAOLT

Trigger	Name the condition that activates MAOLT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-34 - Moussaif Agentic Observability-Ledger Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-35 - Moussaif Dynamic Benchmark-Court Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: static benchmarks becoming targets rather than measurements. Core claim: benchmarks should evolve with hidden, adversarial and transfer cases under court governance. Boundary analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is benchmark docket and promotion verdict. A verbal explanation without this object is not activation of MDBCT. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: MCIT, MECT, MGSET. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MDBCT improve generalization, leakage resistance, drift sensitivity under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MDBCT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets static benchmarks becoming targets rather than measurements. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MDBCT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: moving goalposts and weak reproducibility. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MDBCT

Trigger	Name the condition that activates MDBCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-35 - Moussaif Dynamic Benchmark-Court Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-35 - Moussaif Dynamic Benchmark-Court Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MDBCT transforms an admissible state into benchmark docket and promotion verdict plus an uncertainty and provenance record. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-35 - Moussaif Dynamic Benchmark-Court Theory: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid benchmark docket and promotion verdict. Operational claim analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MDBCT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit benchmark docket and promotion verdict; 7) write telemetry and residual uncertainty. Mechanism specification for TH-35 - Moussaif Dynamic Benchmark-Court Theory: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in generalization, leakage resistance, drift sensitivity exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MCIT, MECT, MGSET. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MDBCT

Trigger	Name the condition that activates MDBCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-35 - Moussaif Dynamic Benchmark-Court Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-35 - Moussaif Dynamic Benchmark-Court Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MDBCT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: benchmark docket and promotion verdict; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-35 - Moussaif Dynamic Benchmark-Court Theory: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MDBCT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure generalization, leakage resistance, drift sensitivity. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-35 - Moussaif Dynamic Benchmark-Court Theory: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by static benchmarks becoming targets rather than measurements, the baseline produces an answer directly. The MDBCT route emits benchmark docket and promotion verdict, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MDBCT

Trigger	Name the condition that activates MDBCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-35 - Moussaif Dynamic Benchmark-Court Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-35 - Moussaif Dynamic Benchmark-Court Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where benchmarks should evolve with hidden, adversarial and transfer cases under court governance. Match models, prompts, budget and tools. If MDBCT does not improve generalization, leakage resistance, drift sensitivity, the claimed mechanism is not supported under that regime. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: moving goalposts and weak reproducibility. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MDBCT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MCIT, MECT, MGSET), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: generalization, leakage resistance, drift sensitivity. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-35 - Moussaif Dynamic Benchmark-Court Theory: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MDBCT

Trigger	Name the condition that activates MDBCT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-35 - Moussaif Dynamic Benchmark-Court Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-36 - Moussaif Governed Self-Evolution Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: self-improvement that changes behavior without proof, rollback or authority. Core claim: every evolution proposal must be staged, benchmarked, approved and reversible. Research-program analysis for TH-36 - Moussaif Governed Self-Evolution Theory: decompose research thesis into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

PRIMARY OPERATIONAL OBJECT

The required artifact is evolution proposal and canary record. A verbal explanation without this object is not activation of MGSET. The object must be schema-valid, versioned and linked to the task contract. Artifact discipline for TH-36 - Moussaif Governed Self-Evolution Theory: the primary operational object claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DEPENDENCIES

Principal dependencies and neighboring constructs: MDBCT, MAOLT, MRRMT. Integration should preserve separate responsibilities and permit component-level ablation. Mechanism specification for TH-36 - Moussaif Governed Self-Evolution Theory: translate dependencies into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

RESEARCH QUESTIONS

Can MGSET improve net gain, regressions, rollback time, policy adherence under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Boundary analysis for TH-36 - Moussaif Governed Self-Evolution Theory: define inclusion, exclusion and abstention regions for research questions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MGSET: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets self-improvement that changes behavior without proof, rollback or authority. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Artifact discipline for TH-36 - Moussaif Governed Self-Evolution Theory: the governing problem claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BOUNDARY AND EXCLUSIONS

MGSET does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Mechanism specification for TH-36 - Moussaif Governed Self-Evolution Theory: translate boundary and exclusions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

KNOWN RISK

Primary conceptual risk: recursive defect amplification and evaluation capture. The registry must record counterexamples, unresolved definitions and alternative explanations. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For known risk, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MGSET

Trigger	Name the condition that activates MGSET; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-36 - Moussaif Governed Self-Evolution Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-36 - Moussaif Governed Self-Evolution Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MGSET transforms an admissible state into evolution proposal and canary record plus an uncertainty and provenance record. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For state and notation, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Artifact discipline for TH-36 - Moussaif Governed Self-Evolution Theory: the invariants claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Integration analysis for TH-36 - Moussaif Governed Self-Evolution Theory: map complexity model to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid evolution proposal and canary record. Boundary analysis for TH-36 - Moussaif Governed Self-Evolution Theory: define inclusion, exclusion and abstention regions for formal failure. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MGSET: outcome المقترح يصبح ذا قيمة فقط عندما يغير baseline وbudget متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit evolution proposal and canary record; 7) write telemetry and residual uncertainty. Integration analysis for TH-36 - Moussaif Governed Self-Evolution Theory: map mechanism to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DECISION RULE

Adopt activation when expected gain in net gain, regressions, rollback time, policy adherence exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Artifact discipline for TH-36 - Moussaif Governed Self-Evolution Theory: the decision rule claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MDBCT, MAOLT, MRRMT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Boundary analysis for TH-36 - Moussaif Governed Self-Evolution Theory: define inclusion, exclusion and abstention regions for interfaces. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MGSET

Trigger	Name the condition that activates MGSET; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-36 - Moussaif Governed Self-Evolution Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-36 - Moussaif Governed Self-Evolution Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MGSET is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Research-program analysis for TH-36 - Moussaif Governed Self-Evolution Theory: decompose runtime architecture into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OUTPUT CONTRACT

Output: evolution proposal and canary record; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Research-program analysis for TH-36 - Moussaif Governed Self-Evolution Theory: decompose output contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Integration analysis for TH-36 - Moussaif Governed Self-Evolution Theory: map security and authority to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For integration verdicts, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MGSET: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Boundary analysis for TH-36 - Moussaif Governed Self-Evolution Theory: define inclusion, exclusion and abstention regions for input contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure net gain, regressions, rollback time, policy adherence. Sensitive content is redacted or referenced, not copied into telemetry. Risk and counterexample analysis for TH-36 - Moussaif Governed Self-Evolution Theory: turn the named limitation in observability into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

WORKED SCENARIO

Given a task dominated by self-improvement that changes behavior without proof, rollback or authority, the baseline produces an answer directly. The MGSET route emits evolution proposal and canary record, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Mechanism specification for TH-36 - Moussaif Governed Self-Evolution Theory: translate worked scenario into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MGSET

Trigger	Name the condition that activates MGSET; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-36 - Moussaif Governed Self-Evolution Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-36 - Moussaif Governed Self-Evolution Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where every evolution proposal must be staged, benchmarked, approved and reversible. Match models, prompts, budget and tools. If MGSET does not improve net gain, regressions, rollback time, policy adherence, the claimed mechanism is not supported under that regime. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary falsifier, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Artifact discipline for TH-36 - Moussaif Governed Self-Evolution Theory: the baselines claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: recursive defect amplification and evaluation capture. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For adversarial cases, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Boundary analysis for TH-36 - Moussaif Governed Self-Evolution Theory: define inclusion, exclusion and abstention regions for next research action. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MGSET: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MDBCT, MAOLT, MRRMT), the court and the memory write separately. Measure both capability loss and cost reduction. Boundary analysis for TH-36 - Moussaif Governed Self-Evolution Theory: define inclusion, exclusion and abstention regions for ablation plan. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

METRICS

Primary metrics: net gain, regressions, rollback time, policy adherence. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For metrics, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Operational claim analysis for TH-36 - Moussaif Governed Self-Evolution Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For evidence ladder, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MGSET

Trigger	Name the condition that activates MGSET; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-36 - Moussaif Governed Self-Evolution Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: skill libraries growing by near-duplicate additions. Core claim: new skills should be synthesized from capability gaps and merged by semantic plus behavioral equivalence. Integration analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: map research thesis to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

PRIMARY OPERATIONAL OBJECT

The required artifact is skill synthesis packet. A verbal explanation without this object is not activation of MSSDT. The object must be schema-valid, versioned and linked to the task contract. Operational claim analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary operational object, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DEPENDENCIES

Principal dependencies and neighboring constructs: MHSG, MCD, MDBCT. Integration should preserve separate responsibilities and permit component-level ablation. Research-program analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: decompose dependencies into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

RESEARCH QUESTIONS

Can MSSDT improve novel coverage, duplicate rate, transfer gain under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Mechanism specification for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: translate research questions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MSSDT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets skill libraries growing by near-duplicate additions. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Operational claim analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For governing problem, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BOUNDARY AND EXCLUSIONS

MSSDT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Research-program analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: decompose boundary and exclusions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

KNOWN RISK

Primary conceptual risk: over-merging distinct edge behavior. The registry must record counterexamples, unresolved definitions and alternative explanations. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in known risk into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MSSDT

Trigger	Name the condition that activates MSSDT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-37 - Moussaif Skill-Synthesis and Deduplication Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MSSDT transforms an admissible state into skill synthesis packet plus an uncertainty and provenance record. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in state and notation into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Operational claim analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For invariants, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Artifact discipline for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: the complexity model claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid skill synthesis packet. Mechanism specification for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: translate formal failure into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MSSDT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit skill synthesis packet; 7) write telemetry and residual uncertainty. Artifact discipline for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: the mechanism claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DECISION RULE

Adopt activation when expected gain in novel coverage, duplicate rate, transfer gain exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Operational claim analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For decision rule, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MHSG, MCD, MDBCT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Mechanism specification for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: translate interfaces into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MSSDT

Trigger	Name the condition that activates MSSDT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-37 - Moussaif Skill-Synthesis and Deduplication Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MSSDT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Integration analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: map runtime architecture to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OUTPUT CONTRACT

Output: skill synthesis packet; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Integration analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: map output contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Artifact discipline for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: the security and authority claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in integration verdicts into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MSSDT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Mechanism specification for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: translate input contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure novel coverage, duplicate rate, transfer gain. Sensitive content is redacted or referenced, not copied into telemetry. Boundary analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: define inclusion, exclusion and abstention regions for observability. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

WORKED SCENARIO

Given a task dominated by skill libraries growing by near-duplicate additions, the baseline produces an answer directly. The MSSDT route emits skill synthesis packet, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Research-program analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: decompose worked scenario into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MSSDT

Trigger	Name the condition that activates MSSDT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-37 - Moussaif Skill-Synthesis and Deduplication Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where new skills should be synthesized from capability gaps and merged by semantic plus behavioral equivalence. Match models, prompts, budget and tools. If MSSDT does not improve novel coverage, duplicate rate, transfer gain, the claimed mechanism is not supported under that regime. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in primary falsifier into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Operational claim analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For baselines, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: over-merging distinct edge behavior. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in adversarial cases into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Mechanism specification for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: translate next research action into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MSSDT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MHSG, MCD, MDBCT), the court and the memory write separately. Measure both capability loss and cost reduction. Mechanism specification for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: translate ablation plan into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

METRICS

Primary metrics: novel coverage, duplicate rate, transfer gain. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in metrics into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Risk and counterexample analysis for TH-37 - Moussaif Skill-Synthesis and Deduplication Theory: turn the named limitation in evidence ladder into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MSSDT

Trigger	Name the condition that activates MSSDT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-37 - Moussaif Skill-Synthesis and Deduplication Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-38 - Moussaif Human-Agency Symbiosis Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: automation displacing human intent, judgment or learning. Core claim: systems should increase human option value, comprehension and revocable control. Mechanism specification for TH-38 - Moussaif Human-Agency Symbiosis Theory: translate research thesis into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

PRIMARY OPERATIONAL OBJECT

The required artifact is human-agency interaction contract. A verbal explanation without this object is not activation of MHAST. The object must be schema-valid, versioned and linked to the task contract. Integration analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: map primary operational object to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

DEPENDENCIES

Principal dependencies and neighboring constructs: MCIL, MABT, MEVE. Integration should preserve separate responsibilities and permit component-level ablation. Boundary analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: define inclusion, exclusion and abstention regions for dependencies. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

RESEARCH QUESTIONS

Can MHAST improve decision quality, control retention, cognitive load, trust calibration under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Risk and counterexample analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: turn the named limitation in research questions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MHAST: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets automation displacing human intent, judgment or learning. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Integration analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: map governing problem to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BOUNDARY AND EXCLUSIONS

MHAST does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Boundary analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: define inclusion, exclusion and abstention regions for boundary and exclusions. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

KNOWN RISK

Primary conceptual risk: automation bias and superficial consent. The registry must record counterexamples, unresolved definitions and alternative explanations. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the known risk claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MHAST

Trigger	Name the condition that activates MHAST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-38 - Moussaif Human-Agency Symbiosis Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-38 - Moussaif Human-Agency Symbiosis Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MHAST transforms an admissible state into human-agency interaction contract plus an uncertainty and provenance record. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the state and notation claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Integration analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: map invariants to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Research-program analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: decompose complexity model into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid human-agency interaction contract. Risk and counterexample analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: turn the named limitation in formal failure into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MHAST: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit human-agency interaction contract; 7) write telemetry and residual uncertainty. Research-program analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: decompose mechanism into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DECISION RULE

Adopt activation when expected gain in decision quality, control retention, cognitive load, trust calibration exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Integration analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: map decision rule to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MCIL, MABT, MEVE. Outputs feed planning, routing, courts, deployment or memory according to the object type. Risk and counterexample analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: turn the named limitation in interfaces into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MHAST

Trigger	Name the condition that activates MHAST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-38 - Moussaif Human-Agency Symbiosis Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-38 - Moussaif Human-Agency Symbiosis Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MHAST is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Mechanism specification for TH-38 - Moussaif Human-Agency Symbiosis Theory: translate runtime architecture into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OUTPUT CONTRACT

Output: human-agency interaction contract; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Mechanism specification for TH-38 - Moussaif Human-Agency Symbiosis Theory: translate output contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Research-program analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: decompose security and authority into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the integration verdicts claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

MHAST: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Risk and counterexample analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: turn the named limitation in input contract into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure decision quality, control retention, cognitive load, trust calibration. Sensitive content is redacted or referenced, not copied into telemetry. Operational claim analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For observability, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

WORKED SCENARIO

Given a task dominated by automation displacing human intent, judgment or learning, the baseline produces an answer directly. The MHAST route emits human-agency interaction contract, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Boundary analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: define inclusion, exclusion and abstention regions for worked scenario. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MHAST

Trigger	Name the condition that activates MHAST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-38 - Moussaif Human-Agency Symbiosis Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-38 - Moussaif Human-Agency Symbiosis Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where systems should increase human option value, comprehension and revocable control. Match models, prompts, budget and tools. If MHAST does not improve decision quality, control retention, cognitive load, trust calibration, the claimed mechanism is not supported under that regime. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the primary falsifier claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Integration analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: map baselines to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: automation bias and superficial consent. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the adversarial cases claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Risk and counterexample analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: turn the named limitation in next research action into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MHAST: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MCIL, MABT, MEVE), the court and the memory write separately. Measure both capability loss and cost reduction. Risk and counterexample analysis for TH-38 - Moussaif Human-Agency Symbiosis Theory: turn the named limitation in ablation plan into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

METRICS

Primary metrics: decision quality, control retention, cognitive load, trust calibration. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the metrics claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Artifact discipline for TH-38 - Moussaif Human-Agency Symbiosis Theory: the evidence ladder claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OPERATIONAL EVIDENCE MATRIX

MHAST

Trigger	Name the condition that activates MHAST; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-38 - Moussaif Human-Agency Symbiosis Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-39 - Moussaif Evidence-to-Value Engine: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: technically correct outputs disconnected from stakeholder value. Core claim: evidence, action and measured world effect should trace to explicit value hypotheses. Integration analysis for TH-39 - Moussaif Evidence-to-Value Engine: map research thesis to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

PRIMARY OPERATIONAL OBJECT

The required artifact is evidence-to-value map. A verbal explanation without this object is not activation of MEVE. The object must be schema-valid, versioned and linked to the task contract. Operational claim analysis for TH-39 - Moussaif Evidence-to-Value Engine: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary operational object, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DEPENDENCIES

Principal dependencies and neighboring constructs: MCD, MPWG, MMT. Integration should preserve separate responsibilities and permit component-level ablation. Research-program analysis for TH-39 - Moussaif Evidence-to-Value Engine: decompose dependencies into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

RESEARCH QUESTIONS

Can MEVE improve realized value, attribution confidence, harm, opportunity cost under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Mechanism specification for TH-39 - Moussaif Evidence-to-Value Engine: translate research questions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MEVE: budget و baseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets technically correct outputs disconnected from stakeholder value. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Operational claim analysis for TH-39 - Moussaif Evidence-to-Value Engine: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For governing problem, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BOUNDARY AND EXCLUSIONS

MEVE does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Research-program analysis for TH-39 - Moussaif Evidence-to-Value Engine: decompose boundary and exclusions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

KNOWN RISK

Primary conceptual risk: proxy gaming and unmeasured externalities. The registry must record counterexamples, unresolved definitions and alternative explanations. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in known risk into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MEVE

Trigger	Name the condition that activates MEVE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-39 - Moussaif Evidence-to-Value Engine.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-39 - Moussaif Evidence-to-Value Engine: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MEVE transforms an admissible state into evidence-to-value map plus an uncertainty and provenance record. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in state and notation into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Operational claim analysis for TH-39 - Moussaif Evidence-to-Value Engine: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For invariants, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Artifact discipline for TH-39 - Moussaif Evidence-to-Value Engine: the complexity model claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid evidence-to-value map. Mechanism specification for TH-39 - Moussaif Evidence-to-Value Engine: translate formal failure into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MEVE: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit evidence-to-value map; 7) write telemetry and residual uncertainty. Artifact discipline for TH-39 - Moussaif Evidence-to-Value Engine: the mechanism claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DECISION RULE

Adopt activation when expected gain in realized value, attribution confidence, harm, opportunity cost exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Operational claim analysis for TH-39 - Moussaif Evidence-to-Value Engine: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For decision rule, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MCD, MPWG, MMT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Mechanism specification for TH-39 - Moussaif Evidence-to-Value Engine: translate interfaces into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MEVE

Trigger	Name the condition that activates MEVE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-39 - Moussaif Evidence-to-Value Engine.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-39 - Moussaif Evidence-to-Value Engine: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MEVE is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Integration analysis for TH-39 - Moussaif Evidence-to-Value Engine: map runtime architecture to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OUTPUT CONTRACT

Output: evidence-to-value map; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Integration analysis for TH-39 - Moussaif Evidence-to-Value Engine: map output contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Artifact discipline for TH-39 - Moussaif Evidence-to-Value Engine: the security and authority claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in integration verdicts into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MEVE: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Mechanism specification for TH-39 - Moussaif Evidence-to-Value Engine: translate input contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure realized value, attribution confidence, harm, opportunity cost. Sensitive content is redacted or referenced, not copied into telemetry. Boundary analysis for TH-39 - Moussaif Evidence-to-Value Engine: define inclusion, exclusion and abstention regions for observability. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

WORKED SCENARIO

Given a task dominated by technically correct outputs disconnected from stakeholder value, the baseline produces an answer directly. The MEVE route emits evidence-to-value map, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Research-program analysis for TH-39 - Moussaif Evidence-to-Value Engine: decompose worked scenario into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MEVE

Trigger	Name the condition that activates MEVE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-39 - Moussaif Evidence-to-Value Engine.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-39 - Moussaif Evidence-to-Value Engine: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where evidence, action and measured world effect should trace to explicit value hypotheses. Match models, prompts, budget and tools. If MEVE does not improve realized value, attribution confidence, harm, opportunity cost, the claimed mechanism is not supported under that regime. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in primary falsifier into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Operational claim analysis for TH-39 - Moussaif Evidence-to-Value Engine: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For baselines, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: proxy gaming and unmeasured externalities. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in adversarial cases into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Mechanism specification for TH-39 - Moussaif Evidence-to-Value Engine: translate next research action into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MEVE: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MCD, MPWG, MMT), the court and the memory write separately. Measure both capability loss and cost reduction. Mechanism specification for TH-39 - Moussaif Evidence-to-Value Engine: translate ablation plan into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

METRICS

Primary metrics: realized value, attribution confidence, harm, opportunity cost. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in metrics into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Risk and counterexample analysis for TH-39 - Moussaif Evidence-to-Value Engine: turn the named limitation in evidence ladder into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MEVE

Trigger	Name the condition that activates MEVE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-39 - Moussaif Evidence-to-Value Engine.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-40 - Moussaif Propagative Overlay-Deployment Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: large replacements that make adoption risky and irreversible. Core claim: capabilities can propagate as overlays with adapters, canaries and measured promotion. Artifact discipline for TH-40 - Moussaif Propagative Overlay-Deployment Theory: the research thesis claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

PRIMARY OPERATIONAL OBJECT

The required artifact is overlay deployment plan. A verbal explanation without this object is not activation of MPODT. The object must be schema-valid, versioned and linked to the task contract. Risk and counterexample analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: turn the named limitation in primary operational object into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

DEPENDENCIES

Principal dependencies and neighboring constructs: M-AGIOS-K, MAOLT, MRRMT. Integration should preserve separate responsibilities and permit component-level ablation. Integration analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: map dependencies to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

RESEARCH QUESTIONS

Can MPODT improve adoption speed, regressions, rollback success, compatibility under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Research-program analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: decompose research questions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MPODT: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets large replacements that make adoption risky and irreversible. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Risk and counterexample analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: turn the named limitation in governing problem into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BOUNDARY AND EXCLUSIONS

MPODT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Integration analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: map boundary and exclusions to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

KNOWN RISK

Primary conceptual risk: permanent transitional complexity. The registry must record counterexamples, unresolved definitions and alternative explanations. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for known risk. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MPODT

Trigger	Name the condition that activates MPODT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-40 - Moussaif Propagative Overlay-Deployment Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-40 - Moussaif Propagative Overlay-Deployment Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MPODT transforms an admissible state into overlay deployment plan plus an uncertainty and provenance record. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for state and notation. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Risk and counterexample analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: turn the named limitation in invariants into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Operational claim analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For complexity model, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid overlay deployment plan. Research-program analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: decompose formal failure into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MPODT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit overlay deployment plan; 7) write telemetry and residual uncertainty. Operational claim analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For mechanism, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DECISION RULE

Adopt activation when expected gain in adoption speed, regressions, rollback success, compatibility exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Risk and counterexample analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: turn the named limitation in decision rule into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and M-AGIOS-K, MAOLT, MRRMT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Research-program analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: decompose interfaces into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MPODT

Trigger	Name the condition that activates MPODT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-40 - Moussaif Propagative Overlay-Deployment Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-40 - Moussaif Propagative Overlay-Deployment Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MPODT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Artifact discipline for TH-40 - Moussaif Propagative Overlay-Deployment Theory: the runtime architecture claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

OUTPUT CONTRACT

Output: overlay deployment plan; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Artifact discipline for TH-40 - Moussaif Propagative Overlay-Deployment Theory: the output contract claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Operational claim analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For security and authority, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for integration verdicts. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

MPODT: يصبح ذا قيمة فقط عندما يغير outcome المقترح تحت baseline وbudget متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Research-program analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: decompose input contract into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure adoption speed, regressions, rollback success, compatibility. Sensitive content is redacted or referenced, not copied into telemetry. Mechanism specification for TH-40 - Moussaif Propagative Overlay-Deployment Theory: translate observability into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

WORKED SCENARIO

Given a task dominated by large replacements that make adoption risky and irreversible, the baseline produces an answer directly. The MPODT route emits overlay deployment plan, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Integration analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: map worked scenario to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MPODT

Trigger	Name the condition that activates MPODT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-40 - Moussaif Propagative Overlay-Deployment Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-40 - Moussaif Propagative Overlay-Deployment Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where capabilities can propagate as overlays with adapters, canaries and measured promotion. Match models, prompts, budget and tools. If MPODT does not improve adoption speed, regressions, rollback success, compatibility, the claimed mechanism is not supported under that regime. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for primary falsifier. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Risk and counterexample analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: turn the named limitation in baselines into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: permanent transitional complexity. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for adversarial cases. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Research-program analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: decompose next research action into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

MPODT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (M-AGIOS-K, MAOLT, MRRMT), the court and the memory write separately. Measure both capability loss and cost reduction. Research-program analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: decompose ablation plan into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

METRICS

Primary metrics: adoption speed, regressions, rollback success, compatibility. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for metrics. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Boundary analysis for TH-40 - Moussaif Propagative Overlay-Deployment Theory: define inclusion, exclusion and abstention regions for evidence ladder. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OPERATIONAL EVIDENCE MATRIX

MPODT

Trigger	Name the condition that activates MPODT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-40 - Moussaif Propagative Overlay-Deployment Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-41 - Moussaif Resilient Recovery-Mesh Theory: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: single-point recovery and fragile state restoration. Core claim: distributed recovery nodes should reconstruct safe state from ledger, snapshots and contracts. Integration analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: map research thesis to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

PRIMARY OPERATIONAL OBJECT

The required artifact is recovery mesh plan. A verbal explanation without this object is not activation of MRRMT. The object must be schema-valid, versioned and linked to the task contract. Operational claim analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For primary operational object, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

DEPENDENCIES

Principal dependencies and neighboring constructs: MAOLT, MSMT, MBRRT. Integration should preserve separate responsibilities and permit component-level ablation. Research-program analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: decompose dependencies into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

RESEARCH QUESTIONS

Can MRRMT improve recovery time, recovery point, integrity, degraded-mode success under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Mechanism specification for TH-41 - Moussaif Resilient Recovery-Mesh Theory: translate research questions into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MRRMT: المقترح يصبح ذا قيمة فقط عندما يغير outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets single-point recovery and fragile state restoration. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Operational claim analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For governing problem, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

BOUNDARY AND EXCLUSIONS

MRRMT does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Research-program analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: decompose boundary and exclusions into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

KNOWN RISK

Primary conceptual risk: correlated backups and replaying corrupted state. The registry must record counterexamples, unresolved definitions and alternative explanations. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in known risk into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MRRMT

Trigger	Name the condition that activates MRRMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-41 - Moussaif Resilient Recovery-Mesh Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-41 - Moussaif Resilient Recovery-Mesh Theory: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MRRMT transforms an admissible state into recovery mesh plan plus an uncertainty and provenance record. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in state and notation into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Operational claim analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For invariants, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Artifact discipline for TH-41 - Moussaif Resilient Recovery-Mesh Theory: the complexity model claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid recovery mesh plan. Mechanism specification for TH-41 - Moussaif Resilient Recovery-Mesh Theory: translate formal failure into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MRRMT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit recovery mesh plan; 7) write telemetry and residual uncertainty. Artifact discipline for TH-41 - Moussaif Resilient Recovery-Mesh Theory: the mechanism claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

DECISION RULE

Adopt activation when expected gain in recovery time, recovery point, integrity, degraded-mode success exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Operational claim analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For decision rule, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and MAOLT, MSMT, MBRRT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Mechanism specification for TH-41 - Moussaif Resilient Recovery-Mesh Theory: translate interfaces into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OPERATIONAL EVIDENCE MATRIX

MRRMT

Trigger	Name the condition that activates MRRMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-41 - Moussaif Resilient Recovery-Mesh Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-41 - Moussaif Resilient Recovery-Mesh Theory: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MRRMT is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Integration analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: map runtime architecture to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OUTPUT CONTRACT

Output: recovery mesh plan; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Integration analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: map output contract to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Artifact discipline for TH-41 - Moussaif Resilient Recovery-Mesh Theory: the security and authority claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in integration verdicts into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

MRRMT: يغير عندما فقط قيمة ذا يصبح المقترح outcome تحت baseline وbudget متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Mechanism specification for TH-41 - Moussaif Resilient Recovery-Mesh Theory: translate input contract into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure recovery time, recovery point, integrity, degraded-mode success. Sensitive content is redacted or referenced, not copied into telemetry. Boundary analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: define inclusion, exclusion and abstention regions for observability. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

WORKED SCENARIO

Given a task dominated by single-point recovery and fragile state restoration, the baseline produces an answer directly. The MRRMT route emits recovery mesh plan, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Research-program analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: decompose worked scenario into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

OPERATIONAL EVIDENCE MATRIX

MRRMT

Trigger	Name the condition that activates MRRMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-41 - Moussaif Resilient Recovery-Mesh Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-41 - Moussaif Resilient Recovery-Mesh Theory: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where distributed recovery nodes should reconstruct safe state from ledger, snapshots and contracts. Match models, prompts, budget and tools. If MRRMT does not improve recovery time, recovery point, integrity, degraded-mode success, the claimed mechanism is not supported under that regime. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in primary falsifier into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Operational claim analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For baselines, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: correlated backups and replaying corrupted state. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in adversarial cases into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Mechanism specification for TH-41 - Moussaif Resilient Recovery-Mesh Theory: translate next research action into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

MRRMT: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (MAOLT, MSMT, MBRRT), the court and the memory write separately. Measure both capability loss and cost reduction. Mechanism specification for TH-41 - Moussaif Resilient Recovery-Mesh Theory: translate ablation plan into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

METRICS

Primary metrics: recovery time, recovery point, integrity, degraded-mode success. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in metrics into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Risk and counterexample analysis for TH-41 - Moussaif Resilient Recovery-Mesh Theory: turn the named limitation in evidence ladder into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MRRMT

Trigger	Name the condition that activates MRRMT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-41 - Moussaif Resilient Recovery-Mesh Theory.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-42 - Moussaif Open/Fattah Intelligence Program: Thesis and Scope

RESEARCH THESIS

This entry is a proposed internal research construct. It becomes scientifically useful only when the predicted causal difference is observable against matched alternatives. Core problem: closed, untestable claims about general or opening intelligence. Core claim: an open program should connect hypotheses, artifacts, tests, negative results and public interfaces. Boundary analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: define inclusion, exclusion and abstention regions for research thesis. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

PRIMARY OPERATIONAL OBJECT

The required artifact is research program map and evidence registry. A verbal explanation without this object is not activation of MOFIP. The object must be schema-valid, versioned and linked to the task contract. Research-program analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: decompose primary operational object into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

DEPENDENCIES

Principal dependencies and neighboring constructs: all kernel theories, MUT, MUG, MOUSSAIXT. Integration should preserve separate responsibilities and permit component-level ablation. Risk and counterexample analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: turn the named limitation in dependencies into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

RESEARCH QUESTIONS

Can MOFIP improve replication, artifact completeness, falsification progress under fixed resources? Does its output predict downstream success? Does removal degrade behavior? Does the mechanism transfer across models, domains and tool environments? Operational claim analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For research questions, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MOFIP: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

GOVERNING PROBLEM

The theory targets closed, untestable claims about general or opening intelligence. The null hypothesis is that an ordinary baseline can solve the same class of tasks with equal quality, cost, authority control and recovery. The alternative predicts a measurable advantage from the named mechanism. Research-program analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: decompose governing problem into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

BOUNDARY AND EXCLUSIONS

MOFIP does not by itself prove AGI, consciousness, safety or autonomy. It is not invoked when the task is trivial, when a simpler control suffices, or when its activation adds no decision-relevant information. Risk and counterexample analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: turn the named limitation in boundary and exclusions into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

KNOWN RISK

Primary conceptual risk: branding outrunning evidence and unresolved acronym lineage. The registry must record counterexamples, unresolved definitions and alternative explanations. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map known risk to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MOFIP

Trigger	Name the condition that activates MOFIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-42 - Moussaif Open/Fattah Intelligence Program.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-42 - Moussaif Open/Fattah Intelligence Program: Formal Model and Mechanism

STATE AND NOTATION

Represent task state as S=(G,W,K,A,E,B), where G is the goal contract, W world state, K knowledge state, A authority, E evidence and B budget. MOFIP transforms an admissible state into research program map and evidence registry plus an uncertainty and provenance record. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map state and notation to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

INVARIANTS

No output may exceed authority, hide missing evidence, merge observed with inferred state, or claim external success without returned state. Every accepted object has provenance, timestamp, version, postcondition and a disable path. Research-program analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: decompose invariants into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

COMPLEXITY MODEL

Total cost includes retrieval, model inference, coordination, tool latency, verification and recovery. Complexity should be reported empirically; asymptotic notation alone cannot describe model or human costs. Mechanism specification for TH-42 - Moussaif Open/Fattah Intelligence Program: translate complexity model into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

FORMAL FAILURE

Failure occurs when the object is invalid, causally inert, non-reproducible, authority-violating or inferior to a matched baseline. A fluent narrative is never a substitute for a valid research program map and evidence registry. Operational claim analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For formal failure, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MOFIP: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

MECHANISM

1) validate trigger; 2) collect the smallest sufficient evidence; 3) construct candidate objects; 4) score causal and operational value; 5) reject invalid candidates; 6) emit research program map and evidence registry; 7) write telemetry and residual uncertainty. Mechanism specification for TH-42 - Moussaif Open/Fattah Intelligence Program: translate mechanism into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

DECISION RULE

Adopt activation when expected gain in replication, artifact completeness, falsification progress exceeds added cost, latency and failure surface. Otherwise use the simpler route. Escalate when uncertainty crosses the risk-sensitive threshold. Research-program analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: decompose decision rule into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

INTERFACES

Inputs arrive from the goal contract, world state, evidence ledger and all kernel theories, MUT, MUG, MOUSSAIXT. Outputs feed planning, routing, courts, deployment or memory according to the object type. Operational claim analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For interfaces, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OPERATIONAL EVIDENCE MATRIX

MOFIP

Trigger	Name the condition that activates MOFIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-42 - Moussaif Open/Fattah Intelligence Program.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-42 - Moussaif Open/Fattah Intelligence Program: Architecture and Tool Interfaces

RUNTIME ARCHITECTURE

MOFIP is reached through the sparse router only after a trigger passes. The runtime loads one primary HyperSkill, zero to three supports, the minimum knowledge set, a model portfolio and typed tools. Boundary analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: define inclusion, exclusion and abstention regions for runtime architecture. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

OUTPUT CONTRACT

Output: research program map and evidence registry; status; evidence grade; metrics; assumptions; residual uncertainty; expected world-state delta; postcondition; rollback or remediation route. Boundary analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: define inclusion, exclusion and abstention regions for output contract. A negative activation case is as important as a positive one because indiscriminate routing can add cost and correlated failure. Document tasks where the theory is irrelevant, where evidence is inadequate and where human judgment remains authoritative. Stress boundaries with stale data, conflicting objectives, rate limits, unavailable tools and partial state. A boundary that cannot be tested is a slogan rather than a governed interface.

SECURITY AND AUTHORITY

Apply least privilege, injection isolation, permission checks, data minimization and human approval for consequential or irreversible actions. Authority is checked again at action time. Mechanism specification for TH-42 - Moussaif Open/Fattah Intelligence Program: translate security and authority into an ordered state transition rather than a descriptive analogy. Identify the trigger, pre-state, transformation, intermediate objects, gate, output and post-state. Each step declares an invariant and a failure signal. The trace must allow an independent reviewer to locate where an incorrect outcome arose. If multiple mechanisms explain the same result, design an intervention that separates them. Operational depth is measured by predictive and diagnostic value, not by the number of architectural labels attached to the theory.

INTEGRATION VERDICTS

ADOPT when the mechanism changes outcome reliably; REVISE when the object is useful but tests fail; MERGE when another theory subsumes it; HOLD when evidence is incomplete; REJECT when no unique causal value appears. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map integration verdicts to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

MOFIP: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

INPUT CONTRACT

Inputs: task contract, current world-state snapshot, authority token, evidence pointers, budget and required output schema. Missing critical fields cause HOLD rather than silent inference. Operational claim analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For input contract, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

OBSERVABILITY

Log trigger decision, route, versions, token and latency cost, tool calls, court verdicts and memory writes. Measure replication, artifact completeness, falsification progress. Sensitive content is redacted or referenced, not copied into telemetry. Artifact discipline for TH-42 - Moussaif Open/Fattah Intelligence Program: the observability claim terminates in a typed, versioned and inspectable object. Required fields include task and source IDs, assumptions, provenance, state timestamp, authority, schema version, confidence, evidence grade, expected world-state delta, postcondition and disable path. The object is rejected if it cannot be replayed or if another component can emit an equivalent object without the named theory. Artifact validity is necessary but not sufficient: downstream decision impact and error reduction must also be measured.

WORKED SCENARIO

Given a task dominated by closed, untestable claims about general or opening intelligence, the baseline produces an answer directly. The MOFIP route emits research program map and evidence registry, exposes assumptions, runs a falsification test and reconciles the returned state before commit. Risk and counterexample analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: turn the named limitation in worked scenario into executable red-team scenarios. Search for false positives, silent omissions, proxy gaming, authority leakage, overconfident language, tool-state divergence and recovery failure. For each counterexample, record whether the response is revision, narrowed scope, new gate, stronger evidence or rejection. Do not rescue the theory by redefining success after observation. Residual risk remains explicit in every promotion record and is compared with the baseline risk, not with an imaginary perfect system.

OPERATIONAL EVIDENCE MATRIX

MOFIP

Trigger	Name the condition that activates MOFIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-42 - Moussaif Open/Fattah Intelligence Program.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TH-42 - Moussaif Open/Fattah Intelligence Program: Falsification and Research Program

PRIMARY FALSIFIER

Construct tasks where an open program should connect hypotheses, artifacts, tests, negative results and public interfaces. Match models, prompts, budget and tools. If MOFIP does not improve replication, artifact completeness, falsification progress, the claimed mechanism is not supported under that regime. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map primary falsifier to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

BASELINES

Compare ordinary prompting, a single-agent pipeline, uniform routing, fixed-model execution, no-memory operation and an expert-designed rule. Include a negative case where activation should be rejected. Research-program analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: decompose baselines into a sequence of specification, prototype, internal evaluation, ablation, transfer and independent replication. Pre-register the central causal claim and primary metric where feasible. Release schemas, benchmark generator, configuration, prompts, tool versions, seeds, negative cases and analysis code. Evidence grade advances only when the required artifact and study exist. Publication language must distinguish concept, runnable mechanism, internally supported result and externally replicated finding.

ADVERSARIAL CASES

Attack the mechanism through stale evidence, conflicting goals, spoofed authority, prompt injection, misleading confidence, tool failure and the known risk: branding outrunning evidence and unresolved acronym lineage. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map adversarial cases to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

NEXT RESEARCH ACTION

Publish the object schema, benchmark generator, matched baseline, ablation matrix and negative results. Seek independent replication before using validated or general-intelligence language. Operational claim analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For next research action, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

MOFIP: budget وbaseline تحت outcome المقترح يصبح ذا قيمة فقط عندما يغير متساويين، مع trace قابل للتكرار.

ABLATION PLAN

Remove the trigger, the operational object, each dependency (all kernel theories, MUT, MUG, MOUSSAIXT), the court and the memory write separately. Measure both capability loss and cost reduction. Operational claim analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: formulate the null hypothesis that a simpler baseline produces the same decision and artifact under equal model, tool, data, time and token conditions. The alternative must predict a direction of change before results are observed. For ablation plan, record dependent and independent variables, confounders, sampling rule and the minimum effect that would justify the added mechanism. Retain null and adverse results. A result tied to one prompt is exploratory; stronger support requires hidden tasks, repeated runs, uncertainty intervals and transfer to a changed model or domain.

METRICS

Primary metrics: replication, artifact completeness, falsification progress. Secondary metrics: calibration, evidence completeness, token cost, latency, route sparsity, recovery time, unauthorized action and human correction. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map metrics to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

EVIDENCE LADDER

E0 idea; E1 specified mechanism; E2 runnable prototype; E3 controlled internal evidence; E4 independent replication; E5 robust multi-domain support. This atlas does not automatically promote any entry. Integration analysis for TH-42 - Moussaif Open/Fattah Intelligence Program: map evidence ladder to upstream producers, downstream consumers, shared schemas and competing components. Every dependency states what it supplies, what freshness and evidence grade are required, and how failure propagates. Remove each dependency separately, then test combined removal to reveal interaction effects. A combination is complementary only when joint performance exceeds the best simpler configuration after accounting for cost, latency and added attack surface. Cycles and hidden global state are treated as architecture defects.

OPERATIONAL EVIDENCE MATRIX

MOFIP

Trigger	Name the condition that activates MOFIP; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TH-42 - Moussaif Open/Fattah Intelligence Program.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-01 - Task Contract Compiler: Typed Contract and Operation

PURPOSE

Task Contract Compiler is a contract and state tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-01 - Task Contract Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-01 - Task Contract Compiler: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Task Contract Compiler result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-01 - Task Contract Compiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-01 - Task Contract Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Task Contract Compiler: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-01 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-01 - Task Contract Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Observability design for TL-01 - Task Contract Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-01 - Task Contract Compiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-01

Trigger	Name the condition that activates TL-01; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-01 - Task Contract Compiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-01 - Task Contract Compiler: Verification, Security and Recovery

VERIFICATION SUITE

Test Task Contract Compiler with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-01 - Task Contract Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-01 - Task Contract Compiler: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-01 - Task Contract Compiler: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-01 - Task Contract Compiler: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Task Contract Compiler: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-01 - Task Contract Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-01, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-01 - Task Contract Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Task Contract Compiler with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-01 - Task Contract Compiler: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-01

Trigger	Name the condition that activates TL-01; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-01 - Task Contract Compiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-02 - Governing Question Generator: Typed Contract and Operation

PURPOSE

Governing Question Generator is a contract and state tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-02 - Governing Question Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-02 - Governing Question Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Governing Question Generator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-02 - Governing Question Generator: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-02 - Governing Question Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Governing Question Generator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط postcondition بعد وreturned state.

INVOCATION TRIGGER

Invoke TL-02 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-02 - Governing Question Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. State-transition detail for TL-02 - Governing Question Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-02 - Governing Question Generator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-02

Trigger	Name the condition that activates TL-02; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-02 - Governing Question Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-02 - Governing Question Generator: Verification, Security and Recovery

VERIFICATION SUITE

Test Governing Question Generator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-02 - Governing Question Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-02 - Governing Question Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-02 - Governing Question Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-02 - Governing Question Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Governing Question Generator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط postcondition بعد وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-02 - Governing Question Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-02, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-02 - Governing Question Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Governing Question Generator with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-02 - Governing Question Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-02

Trigger	Name the condition that activates TL-02; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-02 - Governing Question Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-03 - Assumption Register: Typed Contract and Operation

PURPOSE

Assumption Register is a contract and state tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-03 - Assumption Register: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-03 - Assumption Register: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Assumption Register result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-03 - Assumption Register: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-03 - Assumption Register: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Assumption Register: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-03 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-03 - Assumption Register: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Failure engineering for TL-03 - Assumption Register: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-03 - Assumption Register: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-03

Trigger	Name the condition that activates TL-03; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-03 - Assumption Register.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-03 - Assumption Register: Verification, Security and Recovery

VERIFICATION SUITE

Test Assumption Register with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-03 - Assumption Register: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-03 - Assumption Register: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-03 - Assumption Register: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-03 - Assumption Register: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Assumption Register: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-03 - Assumption Register: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-03, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-03 - Assumption Register: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Assumption Register with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-03 - Assumption Register: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-03

Trigger	Name the condition that activates TL-03; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-03 - Assumption Register.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-04 - Operational Definition Builder: Typed Contract and Operation

PURPOSE

Operational Definition Builder is a contract and state tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-04 - Operational Definition Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-04 - Operational Definition Builder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Operational Definition Builder result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-04 - Operational Definition Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-04 - Operational Definition Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Operational Definition Builder: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-04 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-04 - Operational Definition Builder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Failure engineering for TL-04 - Operational Definition Builder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-04 - Operational Definition Builder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX — TL-04

Trigger	Name the condition that activates TL-04; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-04 - Operational Definition Builder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-04 - Operational Definition Builder: Verification, Security and Recovery

VERIFICATION SUITE

Test Operational Definition Builder with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-04 - Operational Definition Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-04 - Operational Definition Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-04 - Operational Definition Builder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-04 - Operational Definition Builder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Operational Definition Builder: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-04 - Operational Definition Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-04, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-04 - Operational Definition Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Operational Definition Builder with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-04 - Operational Definition Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-04

Trigger	Name the condition that activates TL-04; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-04 - Operational Definition Builder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-05 - Constraint Extractor: Typed Contract and Operation

PURPOSE

Constraint Extractor is a contract and state tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-05 - Constraint Extractor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-05 - Constraint Extractor: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Constraint Extractor result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-05 - Constraint Extractor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-05 - Constraint Extractor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Constraint Extractor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-05 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-05 - Constraint Extractor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Verification design for TL-05 - Constraint Extractor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-05 - Constraint Extractor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-05

Trigger	Name the condition that activates TL-05; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-05 - Constraint Extractor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-05 - Constraint Extractor: Verification, Security and Recovery

VERIFICATION SUITE

Test Constraint Extractor with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-05 - Constraint Extractor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-05 - Constraint Extractor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-05 - Constraint Extractor: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-05 - Constraint Extractor: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Constraint Extractor: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-05 - Constraint Extractor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-05, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-05 - Constraint Extractor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Constraint Extractor with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-05 - Constraint Extractor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-05

Trigger	Name the condition that activates TL-05; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-05 - Constraint Extractor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-06 - Stakeholder Utility Mapper: Typed Contract and Operation

PURPOSE

Stakeholder Utility Mapper is a contract and state tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-06 - Stakeholder Utility Mapper: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-06 - Stakeholder Utility Mapper: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Stakeholder Utility Mapper result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-06 - Stakeholder Utility Mapper: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-06 - Stakeholder Utility Mapper: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Stakeholder Utility Mapper: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-06 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-06 - Stakeholder Utility Mapper: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Lifecycle and integration for TL-06 - Stakeholder Utility Mapper: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-06 - Stakeholder Utility Mapper: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-06

Trigger	Name the condition that activates TL-06; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-06 - Stakeholder Utility Mapper.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-06 - Stakeholder Utility Mapper: Verification, Security and Recovery

VERIFICATION SUITE

Test Stakeholder Utility Mapper with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-06 - Stakeholder Utility Mapper: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-06 - Stakeholder Utility Mapper: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-06 - Stakeholder Utility Mapper: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-06 - Stakeholder Utility Mapper: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Stakeholder Utility Mapper: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-06 - Stakeholder Utility Mapper: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-06, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-06 - Stakeholder Utility Mapper: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Stakeholder Utility Mapper with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-06 - Stakeholder Utility Mapper: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-06

Trigger	Name the condition that activates TL-06; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-06 - Stakeholder Utility Mapper.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-07 - Authority Token Validator: Typed Contract and Operation

PURPOSE

Authority Token Validator is a contract and state tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-07 - Authority Token Validator: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-07 - Authority Token Validator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Authority Token Validator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-07 - Authority Token Validator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-07 - Authority Token Validator: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Authority Token Validator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-07 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-07 - Authority Token Validator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Security analysis for TL-07 - Authority Token Validator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-07 - Authority Token Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-07

Trigger	Name the condition that activates TL-07; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-07 - Authority Token Validator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-07 - Authority Token Validator: Verification, Security and Recovery

VERIFICATION SUITE

Test Authority Token Validator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-07 - Authority Token Validator: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-07 - Authority Token Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-07 - Authority Token Validator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-07 - Authority Token Validator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Authority Token Validator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-07 - Authority Token Validator: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-07, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-07 - Authority Token Validator: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Authority Token Validator with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-07 - Authority Token Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-07

Trigger	Name the condition that activates TL-07; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-07 - Authority Token Validator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-08 - Stop-Condition Compiler: Typed Contract and Operation

PURPOSE

Stop-Condition Compiler is a contract and state tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-08 - Stop-Condition Compiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-08 - Stop-Condition Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Stop-Condition Compiler result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-08 - Stop-Condition Compiler: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-08 - Stop-Condition Compiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Stop-Condition Compiler: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-08 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-08 - Stop-Condition Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Failure engineering for TL-08 - Stop-Condition Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-08 - Stop-Condition Compiler: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-08

Trigger	Name the condition that activates TL-08; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-08 - Stop-Condition Compiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-08 - Stop-Condition Compiler: Verification, Security and Recovery

VERIFICATION SUITE

Test Stop-Condition Compiler with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-08 - Stop-Condition Compiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-08 - Stop-Condition Compiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-08 - Stop-Condition Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-08 - Stop-Condition Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Stop-Condition Compiler: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-08 - Stop-Condition Compiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-08, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-08 - Stop-Condition Compiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Stop-Condition Compiler with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-08 - Stop-Condition Compiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-08

Trigger	Name the condition that activates TL-08; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-08 - Stop-Condition Compiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-09 - World-State Snapshotter: Typed Contract and Operation

PURPOSE

World-State Snapshotter is a contract and state tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-09 - World-State Snapshotter: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-09 - World-State Snapshotter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes World-State Snapshotter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-09 - World-State Snapshotter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-09 - World-State Snapshotter: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

World-State Snapshotter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-09 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-09 - World-State Snapshotter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Security analysis for TL-09 - World-State Snapshotter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-09 - World-State Snapshotter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-09

Trigger	Name the condition that activates TL-09; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-09 - World-State Snapshotter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-09 - World-State Snapshotter: Verification, Security and Recovery

VERIFICATION SUITE

Test World-State Snapshotter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-09 - World-State Snapshotter: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-09 - World-State Snapshotter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-09 - World-State Snapshotter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-09 - World-State Snapshotter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

World-State Snapshotter: مستقلة؛ نجاحها يثبت فقط بعد الأداة لا تملك authority وpostcondition returned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-09 - World-State Snapshotter: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-09, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-09 - World-State Snapshotter: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose World-State Snapshotter with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-09 - World-State Snapshotter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-09

Trigger	Name the condition that activates TL-09; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-09 - World-State Snapshotter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-10 - Project Graph Builder: Typed Contract and Operation

PURPOSE

Project Graph Builder is a contract and state tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-10 - Project Graph Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-10 - Project Graph Builder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes Project Graph Builder result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-10 - Project Graph Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MGCG, Q-MCD, MWSM. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-10 - Project Graph Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Project Graph Builder: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-10 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-10 - Project Graph Builder: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Contract detail for TL-10 - Project Graph Builder: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-10 - Project Graph Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-10

Trigger	Name the condition that activates TL-10; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-10 - Project Graph Builder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-10 - Project Graph Builder: Verification, Security and Recovery

VERIFICATION SUITE

Test Project Graph Builder with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-10 - Project Graph Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-10 - Project Graph Builder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-10 - Project Graph Builder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-10 - Project Graph Builder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Project Graph Builder: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-10 - Project Graph Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-10, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-10 - Project Graph Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose Project Graph Builder with MGCG, Q-MCD, MWSM, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-10 - Project Graph Builder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-10

Trigger	Name the condition that activates TL-10; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-10 - Project Graph Builder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-11 - Claim Ledger: Typed Contract and Operation

PURPOSE

Claim Ledger is a evidence and memory tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-11 - Claim Ledger: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-11 - Claim Ledger: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Claim Ledger result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-11 - Claim Ledger: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-11 - Claim Ledger: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

postcondition مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك :Claim Ledger وstate returned.

INVOCATION TRIGGER

Invoke TL-11 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-11 - Claim Ledger: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. Security analysis for TL-11 - Claim Ledger: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-11 - Claim Ledger: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-11

Trigger	Name the condition that activates TL-11; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-11 - Claim Ledger.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-11 - Claim Ledger: Verification, Security and Recovery

VERIFICATION SUITE

Test Claim Ledger with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-11 - Claim Ledger: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-11 - Claim Ledger: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-11 - Claim Ledger: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-11 - Claim Ledger: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Claim Ledger: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط وstate returned.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-11 - Claim Ledger: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-11, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-11 - Claim Ledger: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Claim Ledger with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-11 - Claim Ledger: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-11

Trigger	Name the condition that activates TL-11; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-11 - Claim Ledger.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-12 - Evidence Pointer Store: Typed Contract and Operation

PURPOSE

Evidence Pointer Store is a evidence and memory tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-12 - Evidence Pointer Store: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-12 - Evidence Pointer Store: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Evidence Pointer Store result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-12 - Evidence Pointer Store: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-12 - Evidence Pointer Store: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Evidence Pointer Store: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-12 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-12 - Evidence Pointer Store: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. Observability design for TL-12 - Evidence Pointer Store: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-12 - Evidence Pointer Store: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-12

Trigger	Name the condition that activates TL-12; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-12 - Evidence Pointer Store.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-12 - Evidence Pointer Store: Verification, Security and Recovery

VERIFICATION SUITE

Test Evidence Pointer Store with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-12 - Evidence Pointer Store: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-12 - Evidence Pointer Store: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-12 - Evidence Pointer Store: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-12 - Evidence Pointer Store: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Evidence Pointer Store: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-12 - Evidence Pointer Store: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-12, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-12 - Evidence Pointer Store: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Evidence Pointer Store with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-12 - Evidence Pointer Store: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-12

Trigger	Name the condition that activates TL-12; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-12 - Evidence Pointer Store.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-13 - Episodic Memory Writer: Typed Contract and Operation

PURPOSE

Episodic Memory Writer is a evidence and memory tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-13 - Episodic Memory Writer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-13 - Episodic Memory Writer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Episodic Memory Writer result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-13 - Episodic Memory Writer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-13 - Episodic Memory Writer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Episodic Memory Writer: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-13 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-13 - Episodic Memory Writer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. Lifecycle and integration for TL-13 - Episodic Memory Writer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-13 - Episodic Memory Writer: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-13

Trigger	Name the condition that activates TL-13; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-13 - Episodic Memory Writer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-13 - Episodic Memory Writer: Verification, Security and Recovery

VERIFICATION SUITE

Test Episodic Memory Writer with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-13 - Episodic Memory Writer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-13 - Episodic Memory Writer: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-13 - Episodic Memory Writer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-13 - Episodic Memory Writer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Episodic Memory Writer: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-13 - Episodic Memory Writer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-13, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-13 - Episodic Memory Writer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Episodic Memory Writer with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-13 - Episodic Memory Writer: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-13

Trigger	Name the condition that activates TL-13; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-13 - Episodic Memory Writer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-14 - Semantic Memory Curator: Typed Contract and Operation

PURPOSE

Semantic Memory Curator is a evidence and memory tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-14 - Semantic Memory Curator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-14 - Semantic Memory Curator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Semantic Memory Curator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-14 - Semantic Memory Curator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-14 - Semantic Memory Curator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Semantic Memory Curator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-14 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-14 - Semantic Memory Curator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Observability design for TL-14 - Semantic Memory Curator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-14 - Semantic Memory Curator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX — TL-14

Trigger	Name the condition that activates TL-14; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-14 - Semantic Memory Curator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-14 - Semantic Memory Curator: Verification, Security and Recovery

VERIFICATION SUITE

Test Semantic Memory Curator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-14 - Semantic Memory Curator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-14 - Semantic Memory Curator: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-14 - Semantic Memory Curator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-14 - Semantic Memory Curator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Semantic Memory Curator: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-14 - Semantic Memory Curator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-14, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-14 - Semantic Memory Curator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Semantic Memory Curator with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-14 - Semantic Memory Curator: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-14

Trigger	Name the condition that activates TL-14; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-14 - Semantic Memory Curator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-15 - Procedural Memory Registry: Typed Contract and Operation

PURPOSE

Procedural Memory Registry is a evidence and memory tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-15 - Procedural Memory Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-15 - Procedural Memory Registry: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Procedural Memory Registry result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-15 - Procedural Memory Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-15 - Procedural Memory Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Procedural Memory Registry: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-15 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-15 - Procedural Memory Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. Observability design for TL-15 - Procedural Memory Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-15 - Procedural Memory Registry: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-15

Trigger	Name the condition that activates TL-15; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-15 - Procedural Memory Registry.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-15 - Procedural Memory Registry: Verification, Security and Recovery

VERIFICATION SUITE

Test Procedural Memory Registry with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-15 - Procedural Memory Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-15 - Procedural Memory Registry: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-15 - Procedural Memory Registry: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-15 - Procedural Memory Registry: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Procedural Memory Registry: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-15 - Procedural Memory Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-15, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-15 - Procedural Memory Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Procedural Memory Registry with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-15 - Procedural Memory Registry: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-15

Trigger	Name the condition that activates TL-15; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-15 - Procedural Memory Registry.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-16 - Context Capsule Compressor: Typed Contract and Operation

PURPOSE

Context Capsule Compressor is a evidence and memory tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-16 - Context Capsule Compressor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-16 - Context Capsule Compressor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Context Capsule Compressor result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-16 - Context Capsule Compressor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-16 - Context Capsule Compressor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Context Capsule Compressor: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-16 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-16 - Context Capsule Compressor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Failure engineering for TL-16 - Context Capsule Compressor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-16 - Context Capsule Compressor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-16

Trigger	Name the condition that activates TL-16; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-16 - Context Capsule Compressor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-16 - Context Capsule Compressor: Verification, Security and Recovery

VERIFICATION SUITE

Test Context Capsule Compressor with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-16 - Context Capsule Compressor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-16 - Context Capsule Compressor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-16 - Context Capsule Compressor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-16 - Context Capsule Compressor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Context Capsule Compressor: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-16 - Context Capsule Compressor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-16, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-16 - Context Capsule Compressor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Context Capsule Compressor with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-16 - Context Capsule Compressor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-16

Trigger	Name the condition that activates TL-16; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-16 - Context Capsule Compressor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-17 - Memory Conflict Resolver: Typed Contract and Operation

PURPOSE

Memory Conflict Resolver is a evidence and memory tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-17 - Memory Conflict Resolver: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-17 - Memory Conflict Resolver: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Memory Conflict Resolver result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-17 - Memory Conflict Resolver: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-17 - Memory Conflict Resolver: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Memory Conflict Resolver: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-17 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-17 - Memory Conflict Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. State-transition detail for TL-17 - Memory Conflict Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-17 - Memory Conflict Resolver: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-17

Trigger	Name the condition that activates TL-17; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-17 - Memory Conflict Resolver.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-17 - Memory Conflict Resolver: Verification, Security and Recovery

VERIFICATION SUITE

Test Memory Conflict Resolver with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-17 - Memory Conflict Resolver: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-17 - Memory Conflict Resolver: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-17 - Memory Conflict Resolver: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-17 - Memory Conflict Resolver: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Memory Conflict Resolver: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-17 - Memory Conflict Resolver: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-17, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-17 - Memory Conflict Resolver: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Memory Conflict Resolver with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-17 - Memory Conflict Resolver: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-17

Trigger	Name the condition that activates TL-17; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-17 - Memory Conflict Resolver.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-18 - Staleness Scanner: Typed Contract and Operation

PURPOSE

Staleness Scanner is a evidence and memory tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-18 - Staleness Scanner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-18 - Staleness Scanner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Staleness Scanner result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-18 - Staleness Scanner: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MELT, MSMT, MCCT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-18 - Staleness Scanner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Staleness Scanner: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-18 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-18 - Staleness Scanner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. State-transition detail for TL-18 - Staleness Scanner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-18 - Staleness Scanner: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-18

Trigger	Name the condition that activates TL-18; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-18 - Staleness Scanner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-18 - Staleness Scanner: Verification, Security and Recovery

VERIFICATION SUITE

Test Staleness Scanner with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-18 - Staleness Scanner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-18 - Staleness Scanner: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-18 - Staleness Scanner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-18 - Staleness Scanner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Staleness Scanner: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-18 - Staleness Scanner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-18, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-18 - Staleness Scanner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Staleness Scanner with MELT, MSMT, MCCT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-18 - Staleness Scanner: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-18

Trigger	Name the condition that activates TL-18; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-18 - Staleness Scanner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-19 - HyperKB Ingestor: Typed Contract and Operation

PURPOSE

HyperKB Ingestor is a knowledge and retrieval tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-19 - HyperKB Ingestor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-19 - HyperKB Ingestor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes HyperKB Ingestor result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-19 - HyperKB Ingestor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-19 - HyperKB Ingestor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

HyperKB Ingestor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-19 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-19 - HyperKB Ingestor: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Contract detail for TL-19 - HyperKB Ingestor: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-19 - HyperKB Ingestor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX — TL-19

Trigger	Name the condition that activates TL-19; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-19 - HyperKB Ingestor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-19 - HyperKB Ingestor: Verification, Security and Recovery

VERIFICATION SUITE

Test HyperKB Ingestor with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-19 - HyperKB Ingestor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-19 - HyperKB Ingestor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-19 - HyperKB Ingestor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-19 - HyperKB Ingestor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

HyperKB Ingestor: مستقلة؛ نجاحها يثبت فقط بعد الأداة لا تملك authority
postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-19 - HyperKB Ingestor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-19, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-19 - HyperKB Ingestor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose HyperKB Ingestor with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-19 - HyperKB Ingestor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-19

Trigger	Name the condition that activates TL-19; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-19 - HyperKB Ingestor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-20 - Chunk Boundary Parser: Typed Contract and Operation

PURPOSE

Chunk Boundary Parser is a knowledge and retrieval tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-20 - Chunk Boundary Parser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-20 - Chunk Boundary Parser: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Chunk Boundary Parser result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-20 - Chunk Boundary Parser: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-20 - Chunk Boundary Parser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Chunk Boundary Parser: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-20 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-20 - Chunk Boundary Parser: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Lifecycle and integration for TL-20 - Chunk Boundary Parser: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-20 - Chunk Boundary Parser: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-20

Trigger	Name the condition that activates TL-20; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-20 - Chunk Boundary Parser.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-20 - Chunk Boundary Parser: Verification, Security and Recovery

VERIFICATION SUITE

Test Chunk Boundary Parser with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-20 - Chunk Boundary Parser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-20 - Chunk Boundary Parser: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-20 - Chunk Boundary Parser: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-20 - Chunk Boundary Parser: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Chunk Boundary Parser: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-20 - Chunk Boundary Parser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-20, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-20 - Chunk Boundary Parser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Chunk Boundary Parser with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-20 - Chunk Boundary Parser: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-20

Trigger	Name the condition that activates TL-20; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-20 - Chunk Boundary Parser.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-21 - Hybrid Retriever: Typed Contract and Operation

PURPOSE

Hybrid Retriever is a knowledge and retrieval tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-21 - Hybrid Retriever: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-21 - Hybrid Retriever: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Hybrid Retriever result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-21 - Hybrid Retriever: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-21 - Hybrid Retriever: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Hybrid Retriever: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-21 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-21 - Hybrid Retriever: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Failure engineering for TL-21 - Hybrid Retriever: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-21 - Hybrid Retriever: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX — TL-21

Trigger	Name the condition that activates TL-21; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-21 - Hybrid Retriever.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-21 - Hybrid Retriever: Verification, Security and Recovery

VERIFICATION SUITE

Test Hybrid Retriever with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-21 - Hybrid Retriever: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-21 - Hybrid Retriever: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-21 - Hybrid Retriever: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-21 - Hybrid Retriever: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Hybrid Retriever: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-21 - Hybrid Retriever: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-21, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-21 - Hybrid Retriever: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Hybrid Retriever with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-21 - Hybrid Retriever: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-21

Trigger	Name the condition that activates TL-21; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-21 - Hybrid Retriever.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-22 - GraphRAG Traverser: Typed Contract and Operation

PURPOSE

GraphRAG Traverser is a knowledge and retrieval tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-22 - GraphRAG Traverser: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-22 - GraphRAG Traverser: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes GraphRAG Traverser result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-22 - GraphRAG Traverser: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-22 - GraphRAG Traverser: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

GraphRAG Traverser: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-22 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-22 - GraphRAG Traverser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Security analysis for TL-22 - GraphRAG Traverser: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-22 - GraphRAG Traverser: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-22

Trigger	Name the condition that activates TL-22; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-22 - GraphRAG Traverser.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-22 - GraphRAG Traverser: Verification, Security and Recovery

VERIFICATION SUITE

Test GraphRAG Traverser with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-22 - GraphRAG Traverser: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-22 - GraphRAG Traverser: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-22 - GraphRAG Traverser: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-22 - GraphRAG Traverser: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

GraphRAG Traverser: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-22 - GraphRAG Traverser: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-22, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-22 - GraphRAG Traverser: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose GraphRAG Traverser with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-22 - GraphRAG Traverser: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-22

Trigger	Name the condition that activates TL-22; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-22 - GraphRAG Traverser.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-23 - Provenance Filter: Typed Contract and Operation

PURPOSE

Provenance Filter is a knowledge and retrieval tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-23 - Provenance Filter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-23 - Provenance Filter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Provenance Filter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-23 - Provenance Filter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-23 - Provenance Filter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Provenance Filter: مستقلة؛ نجاحها يثبت فقط بعد الأداة لا تملك authority returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-23 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-23 - Provenance Filter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Lifecycle and integration for TL-23 - Provenance Filter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-23 - Provenance Filter: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-23

Trigger	Name the condition that activates TL-23; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-23 - Provenance Filter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-23 - Provenance Filter: Verification, Security and Recovery

VERIFICATION SUITE

Test Provenance Filter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-23 - Provenance Filter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-23 - Provenance Filter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-23 - Provenance Filter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-23 - Provenance Filter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Provenance Filter: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-23 - Provenance Filter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-23, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-23 - Provenance Filter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Provenance Filter with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-23 - Provenance Filter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-23

Trigger	Name the condition that activates TL-23; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-23 - Provenance Filter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-24 - Freshness Gate: Typed Contract and Operation

PURPOSE

Freshness Gate is a knowledge and retrieval tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-24 - Freshness Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-24 - Freshness Gate: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Freshness Gate result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-24 - Freshness Gate: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-24 - Freshness Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Freshness Gate: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وحالته.

INVOCATION TRIGGER

Invoke TL-24 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-24 - Freshness Gate: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Verification design for TL-24 - Freshness Gate: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-24 - Freshness Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-24

Trigger	Name the condition that activates TL-24; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-24 - Freshness Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-24 - Freshness Gate: Verification, Security and Recovery

VERIFICATION SUITE

Test Freshness Gate with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-24 - Freshness Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-24 - Freshness Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-24 - Freshness Gate: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-24 - Freshness Gate: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Freshness Gate: postcondition مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك وstate returned.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-24 - Freshness Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-24, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-24 - Freshness Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Freshness Gate with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-24 - Freshness Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-24

Trigger	Name the condition that activates TL-24; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-24 - Freshness Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-25 - Citation Verifier: Typed Contract and Operation

PURPOSE

Citation Verifier is a knowledge and retrieval tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-25 - Citation Verifier: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-25 - Citation Verifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Citation Verifier result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-25 - Citation Verifier: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-25 - Citation Verifier: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Citation Verifier: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وstate.

INVOCATION TRIGGER

Invoke TL-25 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-25 - Citation Verifier: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. State-transition detail for TL-25 - Citation Verifier: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-25 - Citation Verifier: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-25

Trigger	Name the condition that activates TL-25; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-25 - Citation Verifier.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-25 - Citation Verifier: Verification, Security and Recovery

VERIFICATION SUITE

Test Citation Verifier with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-25 - Citation Verifier: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-25 - Citation Verifier: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-25 - Citation Verifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-25 - Citation Verifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Citation Verifier: postcondition مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-25 - Citation Verifier: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-25, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-25 - Citation Verifier: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Citation Verifier with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-25 - Citation Verifier: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-25

Trigger	Name the condition that activates TL-25; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-25 - Citation Verifier.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-26 - Retrieval Quality Evaluator: Typed Contract and Operation

PURPOSE

Retrieval Quality Evaluator is a knowledge and retrieval tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-26 - Retrieval Quality Evaluator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-26 - Retrieval Quality Evaluator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Retrieval Quality Evaluator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-26 - Retrieval Quality Evaluator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-26 - Retrieval Quality Evaluator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Retrieval Quality Evaluator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-26 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-26 - Retrieval Quality Evaluator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. Observability design for TL-26 - Retrieval Quality Evaluator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-26 - Retrieval Quality Evaluator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-26

Trigger	Name the condition that activates TL-26; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-26 - Retrieval Quality Evaluator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-26 - Retrieval Quality Evaluator: Verification, Security and Recovery

VERIFICATION SUITE

Test Retrieval Quality Evaluator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-26 - Retrieval Quality Evaluator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-26 - Retrieval Quality Evaluator: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-26 - Retrieval Quality Evaluator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-26 - Retrieval Quality Evaluator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Retrieval Quality Evaluator: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-26 - Retrieval Quality Evaluator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-26, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-26 - Retrieval Quality Evaluator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Retrieval Quality Evaluator with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-26 - Retrieval Quality Evaluator: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-26

Trigger	Name the condition that activates TL-26; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-26 - Retrieval Quality Evaluator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-27 - Knowledge Gap Detector: Typed Contract and Operation

PURPOSE

Knowledge Gap Detector is a knowledge and retrieval tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-27 - Knowledge Gap Detector: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-27 - Knowledge Gap Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Knowledge Gap Detector result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-27 - Knowledge Gap Detector: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-27 - Knowledge Gap Detector: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Knowledge Gap Detector: الأداة لا تملك authority بعد فقط يثبت مستقلة؛ نجاحها returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-27 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-27 - Knowledge Gap Detector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. State-transition detail for TL-27 - Knowledge Gap Detector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-27 - Knowledge Gap Detector: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-27

Trigger	Name the condition that activates TL-27; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-27 - Knowledge Gap Detector.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-27 - Knowledge Gap Detector: Verification, Security and Recovery

VERIFICATION SUITE

Test Knowledge Gap Detector with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-27 - Knowledge Gap Detector: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-27 - Knowledge Gap Detector: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-27 - Knowledge Gap Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-27 - Knowledge Gap Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Knowledge Gap Detector: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned stateو postcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-27 - Knowledge Gap Detector: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-27, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-27 - Knowledge Gap Detector: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Knowledge Gap Detector with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-27 - Knowledge Gap Detector: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-27

Trigger	Name the condition that activates TL-27; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-27 - Knowledge Gap Detector.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-28 - Canonical Memory Promoter: Typed Contract and Operation

PURPOSE

Canonical Memory Promoter is a knowledge and retrieval tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-28 - Canonical Memory Promoter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-28 - Canonical Memory Promoter: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Canonical Memory Promoter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-28 - Canonical Memory Promoter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MHKOS, MPART, MELT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-28 - Canonical Memory Promoter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Canonical Memory Promoter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-28 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-28 - Canonical Memory Promoter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Verification design for TL-28 - Canonical Memory Promoter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-28 - Canonical Memory Promoter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-28

Trigger	Name the condition that activates TL-28; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-28 - Canonical Memory Promoter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-28 - Canonical Memory Promoter: Verification, Security and Recovery

VERIFICATION SUITE

Test Canonical Memory Promoter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-28 - Canonical Memory Promoter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-28 - Canonical Memory Promoter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-28 - Canonical Memory Promoter: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-28 - Canonical Memory Promoter: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Canonical Memory Promoter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-28 - Canonical Memory Promoter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-28, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-28 - Canonical Memory Promoter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Canonical Memory Promoter with MHKOS, MPART, MELT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-28 - Canonical Memory Promoter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-28

Trigger	Name the condition that activates TL-28; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-28 - Canonical Memory Promoter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-29 - MCD Scorer: Typed Contract and Operation

PURPOSE

MCD Scorer is a reasoning and planning tool whose responsibility is to rank candidates using task, evidence, cost, risk and authority signals. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-29 - MCD Scorer: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-29 - MCD Scorer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes MCD Scorer result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-29 - MCD Scorer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-29 - MCD Scorer: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

MCD Scorer: postcondition مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك وstate returned.

INVOCATION TRIGGER

Invoke TL-29 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-29 - MCD Scorer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to rank candidates using task, evidence, cost, risk and authority signals. Security analysis for TL-29 - MCD Scorer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-29 - MCD Scorer: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX — TL-29

Trigger	Name the condition that activates TL-29; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-29 - MCD Scorer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-29 - MCD Scorer: Verification, Security and Recovery

VERIFICATION SUITE

Test MCD Scorer with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-29 - MCD Scorer: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-29 - MCD Scorer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-29 - MCD Scorer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-29 - MCD Scorer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

MCD Scorer: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وحالة.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-29 - MCD Scorer: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-29, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-29 - MCD Scorer: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose MCD Scorer with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-29 - MCD Scorer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-29

Trigger	Name the condition that activates TL-29; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-29 - MCD Scorer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-30 - Q-MCD Gate: Typed Contract and Operation

PURPOSE

Q-MCD Gate is a reasoning and planning tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-30 - Q-MCD Gate: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-30 - Q-MCD Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Q-MCD Gate result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-30 - Q-MCD Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-30 - Q-MCD Gate: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Q-MCD Gate: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وحالة.

INVOCATION TRIGGER

Invoke TL-30 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-30 - Q-MCD Gate: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Security analysis for TL-30 - Q-MCD Gate: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-30 - Q-MCD Gate: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-30

Trigger	Name the condition that activates TL-30; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-30 - Q-MCD Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-30 - Q-MCD Gate: Verification, Security and Recovery

VERIFICATION SUITE

Test Q-MCD Gate with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-30 - Q-MCD Gate: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-30 - Q-MCD Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-30 - Q-MCD Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-30 - Q-MCD Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Q-MCD Gate: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وحالة.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-30 - Q-MCD Gate: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-30, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-30 - Q-MCD Gate: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Q-MCD Gate with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-30 - Q-MCD Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-30

Trigger	Name the condition that activates TL-30; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-30 - Q-MCD Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-31 - Causal Graph Constructor: Typed Contract and Operation

PURPOSE

Causal Graph Constructor is a reasoning and planning tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-31 - Causal Graph Constructor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-31 - Causal Graph Constructor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Causal Graph Constructor result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-31 - Causal Graph Constructor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-31 - Causal Graph Constructor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Causal Graph Constructor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-31 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-31 - Causal Graph Constructor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Lifecycle and integration for TL-31 - Causal Graph Constructor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-31 - Causal Graph Constructor: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-31

Trigger	Name the condition that activates TL-31; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-31 - Causal Graph Constructor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-31 - Causal Graph Constructor: Verification, Security and Recovery

VERIFICATION SUITE

Test Causal Graph Constructor with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-31 - Causal Graph Constructor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-31 - Causal Graph Constructor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-31 - Causal Graph Constructor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-31 - Causal Graph Constructor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Causal Graph Constructor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-31 - Causal Graph Constructor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-31, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-31 - Causal Graph Constructor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Causal Graph Constructor with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-31 - Causal Graph Constructor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-31

Trigger	Name the condition that activates TL-31; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-31 - Causal Graph Constructor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-32 - Hypothesis Generator: Typed Contract and Operation

PURPOSE

Hypothesis Generator is a reasoning and planning tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-32 - Hypothesis Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-32 - Hypothesis Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Hypothesis Generator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-32 - Hypothesis Generator: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-32 - Hypothesis Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Hypothesis Generator: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-32 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-32 - Hypothesis Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. State-transition detail for TL-32 - Hypothesis Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-32 - Hypothesis Generator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-32

Trigger	Name the condition that activates TL-32; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-32 - Hypothesis Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-32 - Hypothesis Generator: Verification, Security and Recovery

VERIFICATION SUITE

Test Hypothesis Generator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-32 - Hypothesis Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-32 - Hypothesis Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-32 - Hypothesis Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-32 - Hypothesis Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Hypothesis Generator: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-32 - Hypothesis Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-32, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-32 - Hypothesis Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Hypothesis Generator with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-32 - Hypothesis Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-32

Trigger	Name the condition that activates TL-32; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-32 - Hypothesis Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-33 - Counterexample Generator: Typed Contract and Operation

PURPOSE

Counterexample Generator is a reasoning and planning tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-33 - Counterexample Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-33 - Counterexample Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Counterexample Generator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-33 - Counterexample Generator: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-33 - Counterexample Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Counterexample Generator: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-33 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-33 - Counterexample Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. State-transition detail for TL-33 - Counterexample Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-33 - Counterexample Generator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-33

Trigger	Name the condition that activates TL-33; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-33 - Counterexample Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-33 - Counterexample Generator: Verification, Security and Recovery

VERIFICATION SUITE

Test Counterexample Generator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-33 - Counterexample Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-33 - Counterexample Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-33 - Counterexample Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-33 - Counterexample Generator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Counterexample Generator: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-33 - Counterexample Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-33, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-33 - Counterexample Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Counterexample Generator with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-33 - Counterexample Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-33

Trigger	Name the condition that activates TL-33; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-33 - Counterexample Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-34 - Abstraction Ladder: Typed Contract and Operation

PURPOSE

Abstraction Ladder is a reasoning and planning tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-34 - Abstraction Ladder: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-34 - Abstraction Ladder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Abstraction Ladder result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-34 - Abstraction Ladder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-34 - Abstraction Ladder: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Abstraction Ladder: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-34 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-34 - Abstraction Ladder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Security analysis for TL-34 - Abstraction Ladder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-34 - Abstraction Ladder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-34

Trigger	Name the condition that activates TL-34; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-34 - Abstraction Ladder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-34 - Abstraction Ladder: Verification, Security and Recovery

VERIFICATION SUITE

Test Abstraction Ladder with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-34 - Abstraction Ladder: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-34 - Abstraction Ladder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-34 - Abstraction Ladder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-34 - Abstraction Ladder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Abstraction Ladder: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-34 - Abstraction Ladder: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-34, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-34 - Abstraction Ladder: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Abstraction Ladder with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-34 - Abstraction Ladder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-34

Trigger	Name the condition that activates TL-34; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-34 - Abstraction Ladder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-35 - Temporal Plan Graph Compiler: Typed Contract and Operation

PURPOSE

Temporal Plan Graph Compiler is a reasoning and planning tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-35 - Temporal Plan Graph Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-35 - Temporal Plan Graph Compiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Temporal Plan Graph Compiler result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-35 - Temporal Plan Graph Compiler: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-35 - Temporal Plan Graph Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Temporal Plan Graph Compiler: مستقلة؛ نجاحها يثبت فقط authority الأداة لا تملك بعد postcondition وreturned state.

INVOCATION TRIGGER

Invoke TL-35 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-35 - Temporal Plan Graph Compiler: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. State-transition detail for TL-35 - Temporal Plan Graph Compiler: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-35 - Temporal Plan Graph Compiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-35

Trigger	Name the condition that activates TL-35; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-35 - Temporal Plan Graph Compiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-35 - Temporal Plan Graph Compiler: Verification, Security and Recovery

VERIFICATION SUITE

Test Temporal Plan Graph Compiler with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-35 - Temporal Plan Graph Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-35 - Temporal Plan Graph Compiler: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-35 - Temporal Plan Graph Compiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-35 - Temporal Plan Graph Compiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Temporal Plan Graph Compiler: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط postcondition بعد وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-35 - Temporal Plan Graph Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-35, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-35 - Temporal Plan Graph Compiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Temporal Plan Graph Compiler with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-35 - Temporal Plan Graph Compiler: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-35

Trigger	Name the condition that activates TL-35; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-35 - Temporal Plan Graph Compiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-36 - Dependency Resolver: Typed Contract and Operation

PURPOSE

Dependency Resolver is a reasoning and planning tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-36 - Dependency Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-36 - Dependency Resolver: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Dependency Resolver result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-36 - Dependency Resolver: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-36 - Dependency Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Dependency Resolver: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-36 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-36 - Dependency Resolver: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Verification design for TL-36 - Dependency Resolver: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-36 - Dependency Resolver: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-36

Trigger	Name the condition that activates TL-36; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-36 - Dependency Resolver.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-36 - Dependency Resolver: Verification, Security and Recovery

VERIFICATION SUITE

Test Dependency Resolver with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-36 - Dependency Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-36 - Dependency Resolver: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-36 - Dependency Resolver: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-36 - Dependency Resolver: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Dependency Resolver: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-36 - Dependency Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-36, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-36 - Dependency Resolver: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Dependency Resolver with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-36 - Dependency Resolver: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-36

Trigger	Name the condition that activates TL-36; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-36 - Dependency Resolver.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-37 - Reversibility Analyzer: Typed Contract and Operation

PURPOSE

Reversibility Analyzer is a reasoning and planning tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-37 - Reversibility Analyzer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-37 - Reversibility Analyzer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Reversibility Analyzer result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-37 - Reversibility Analyzer: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-37 - Reversibility Analyzer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Reversibility Analyzer: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-37 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-37 - Reversibility Analyzer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. Failure engineering for TL-37 - Reversibility Analyzer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-37 - Reversibility Analyzer: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-37

Trigger	Name the condition that activates TL-37; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-37 - Reversibility Analyzer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-37 - Reversibility Analyzer: Verification, Security and Recovery

VERIFICATION SUITE

Test Reversibility Analyzer with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-37 - Reversibility Analyzer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-37 - Reversibility Analyzer: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-37 - Reversibility Analyzer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-37 - Reversibility Analyzer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Reversibility Analyzer: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-37 - Reversibility Analyzer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-37, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-37 - Reversibility Analyzer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Reversibility Analyzer with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-37 - Reversibility Analyzer: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-37

Trigger	Name the condition that activates TL-37; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-37 - Reversibility Analyzer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-38 - Uncertainty Propagator: Typed Contract and Operation

PURPOSE

Uncertainty Propagator is a reasoning and planning tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-38 - Uncertainty Propagator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-38 - Uncertainty Propagator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Uncertainty Propagator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-38 - Uncertainty Propagator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-38 - Uncertainty Propagator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Uncertainty Propagator: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-38 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-38 - Uncertainty Propagator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Observability design for TL-38 - Uncertainty Propagator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-38 - Uncertainty Propagator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-38

Trigger	Name the condition that activates TL-38; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-38 - Uncertainty Propagator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-38 - Uncertainty Propagator: Verification, Security and Recovery

VERIFICATION SUITE

Test Uncertainty Propagator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-38 - Uncertainty Propagator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-38 - Uncertainty Propagator: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-38 - Uncertainty Propagator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-38 - Uncertainty Propagator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Uncertainty Propagator: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-38 - Uncertainty Propagator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-38, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-38 - Uncertainty Propagator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Uncertainty Propagator with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-38 - Uncertainty Propagator: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-38

Trigger	Name the condition that activates TL-38; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-38 - Uncertainty Propagator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-39 - Decision Packet Builder: Typed Contract and Operation

PURPOSE

Decision Packet Builder is a reasoning and planning tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-39 - Decision Packet Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-39 - Decision Packet Builder: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Decision Packet Builder result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-39 - Decision Packet Builder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-39 - Decision Packet Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Decision Packet Builder: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-39 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-39 - Decision Packet Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Verification design for TL-39 - Decision Packet Builder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-39 - Decision Packet Builder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-39

Trigger	Name the condition that activates TL-39; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-39 - Decision Packet Builder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-39 - Decision Packet Builder: Verification, Security and Recovery

VERIFICATION SUITE

Test Decision Packet Builder with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-39 - Decision Packet Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-39 - Decision Packet Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-39 - Decision Packet Builder: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-39 - Decision Packet Builder: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Decision Packet Builder: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-39 - Decision Packet Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-39, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-39 - Decision Packet Builder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Decision Packet Builder with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-39 - Decision Packet Builder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-39

Trigger	Name the condition that activates TL-39; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-39 - Decision Packet Builder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-40 - Action Proposal Ranker: Typed Contract and Operation

PURPOSE

Action Proposal Ranker is a reasoning and planning tool whose responsibility is to rank candidates using task, evidence, cost, risk and authority signals. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-40 - Action Proposal Ranker: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-40 - Action Proposal Ranker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Action Proposal Ranker result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-40 - Action Proposal Ranker: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MCD, MTPG, MAECT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-40 - Action Proposal Ranker: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Action Proposal Ranker: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-40 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-40 - Action Proposal Ranker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to rank candidates using task, evidence, cost, risk and authority signals. Lifecycle and integration for TL-40 - Action Proposal Ranker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-40 - Action Proposal Ranker: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-40

Trigger	Name the condition that activates TL-40; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-40 - Action Proposal Ranker.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-40 - Action Proposal Ranker: Verification, Security and Recovery

VERIFICATION SUITE

Test Action Proposal Ranker with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-40 - Action Proposal Ranker: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-40 - Action Proposal Ranker: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-40 - Action Proposal Ranker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-40 - Action Proposal Ranker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Action Proposal Ranker: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-40 - Action Proposal Ranker: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-40, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-40 - Action Proposal Ranker: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Action Proposal Ranker with MCD, MTPG, MAECT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-40 - Action Proposal Ranker: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-40

Trigger	Name the condition that activates TL-40; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-40 - Action Proposal Ranker.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-41 - HyperSkill Registry: Typed Contract and Operation

PURPOSE

HyperSkill Registry is a skills and orchestration tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-41 - HyperSkill Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-41 - HyperSkill Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes HyperSkill Registry result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-41 - HyperSkill Registry: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-41 - HyperSkill Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

HyperSkill Registry: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-41 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-41 - HyperSkill Registry: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. State-transition detail for TL-41 - HyperSkill Registry: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-41 - HyperSkill Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-41

Trigger	Name the condition that activates TL-41; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-41 - HyperSkill Registry.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-41 - HyperSkill Registry: Verification, Security and Recovery

VERIFICATION SUITE

Test HyperSkill Registry with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-41 - HyperSkill Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-41 - HyperSkill Registry: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-41 - HyperSkill Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-41 - HyperSkill Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

HyperSkill Registry: الأداة لا تملك authority بعد فقط يثبت نجاحها؛ مستقلة returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-41 - HyperSkill Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-41, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-41 - HyperSkill Registry: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose HyperSkill Registry with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-41 - HyperSkill Registry: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-41

Trigger	Name the condition that activates TL-41; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-41 - HyperSkill Registry.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-42 - Skill Similarity Deduplicator: Typed Contract and Operation

PURPOSE

Skill Similarity Deduplicator is a skills and orchestration tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-42 - Skill Similarity Deduplicator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-42 - Skill Similarity Deduplicator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Skill Similarity Deduplicator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-42 - Skill Similarity Deduplicator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-42 - Skill Similarity Deduplicator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Skill Similarity Deduplicator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-42 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-42 - Skill Similarity Deduplicator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Failure engineering for TL-42 - Skill Similarity Deduplicator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-42 - Skill Similarity Deduplicator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-42

Trigger	Name the condition that activates TL-42; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-42 - Skill Similarity Deduplicator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-42 - Skill Similarity Deduplicator: Verification, Security and Recovery

VERIFICATION SUITE

Test Skill Similarity Deduplicator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-42 - Skill Similarity Deduplicator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-42 - Skill Similarity Deduplicator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-42 - Skill Similarity Deduplicator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-42 - Skill Similarity Deduplicator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Skill Similarity Deduplicator: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-42 - Skill Similarity Deduplicator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-42, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-42 - Skill Similarity Deduplicator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Skill Similarity Deduplicator with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-42 - Skill Similarity Deduplicator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-42

Trigger	Name the condition that activates TL-42; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-42 - Skill Similarity Deduplicator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-43 - Skill Activation Kernel: Typed Contract and Operation

PURPOSE

Skill Activation Kernel is a skills and orchestration tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-43 - Skill Activation Kernel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-43 - Skill Activation Kernel: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Skill Activation Kernel result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-43 - Skill Activation Kernel: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-43 - Skill Activation Kernel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Skill Activation Kernel: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-43 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-43 - Skill Activation Kernel: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. State-transition detail for TL-43 - Skill Activation Kernel: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-43 - Skill Activation Kernel: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-43

Trigger	Name the condition that activates TL-43; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-43 - Skill Activation Kernel.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-43 - Skill Activation Kernel: Verification, Security and Recovery

VERIFICATION SUITE

Test Skill Activation Kernel with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-43 - Skill Activation Kernel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-43 - Skill Activation Kernel: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-43 - Skill Activation Kernel: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-43 - Skill Activation Kernel: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Skill Activation Kernel: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-43 - Skill Activation Kernel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-43, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-43 - Skill Activation Kernel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Skill Activation Kernel with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-43 - Skill Activation Kernel: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-43

Trigger	Name the condition that activates TL-43; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-43 - Skill Activation Kernel.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-44 - Sparse HyperRouter: Typed Contract and Operation

PURPOSE

Sparse HyperRouter is a skills and orchestration tool whose responsibility is to rank candidates using task, evidence, cost, risk and authority signals. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-44 - Sparse HyperRouter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-44 - Sparse HyperRouter: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Sparse HyperRouter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-44 - Sparse HyperRouter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-44 - Sparse HyperRouter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Sparse HyperRouter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-44 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-44 - Sparse HyperRouter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to rank candidates using task, evidence, cost, risk and authority signals. Verification design for TL-44 - Sparse HyperRouter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-44 - Sparse HyperRouter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-44

Trigger	Name the condition that activates TL-44; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-44 - Sparse HyperRouter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-44 - Sparse HyperRouter: Verification, Security and Recovery

VERIFICATION SUITE

Test Sparse HyperRouter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-44 - Sparse HyperRouter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-44 - Sparse HyperRouter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-44 - Sparse HyperRouter: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-44 - Sparse HyperRouter: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Sparse HyperRouter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-44 - Sparse HyperRouter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-44, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-44 - Sparse HyperRouter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Sparse HyperRouter with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-44 - Sparse HyperRouter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-44

Trigger	Name the condition that activates TL-44; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-44 - Sparse HyperRouter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-45 - Model Capability Profiler: Typed Contract and Operation

PURPOSE

Model Capability Profiler is a skills and orchestration tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-45 - Model Capability Profiler: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-45 - Model Capability Profiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Model Capability Profiler result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-45 - Model Capability Profiler: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-45 - Model Capability Profiler: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Model Capability Profiler: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-45 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-45 - Model Capability Profiler: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. Security analysis for TL-45 - Model Capability Profiler: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-45 - Model Capability Profiler: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-45

Trigger	Name the condition that activates TL-45; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-45 - Model Capability Profiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-45 - Model Capability Profiler: Verification, Security and Recovery

VERIFICATION SUITE

Test Model Capability Profiler with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-45 - Model Capability Profiler: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-45 - Model Capability Profiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-45 - Model Capability Profiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-45 - Model Capability Profiler: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Model Capability Profiler: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-45 - Model Capability Profiler: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-45, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-45 - Model Capability Profiler: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Model Capability Profiler with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-45 - Model Capability Profiler: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-45

Trigger	Name the condition that activates TL-45; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-45 - Model Capability Profiler.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-46 - Model Portfolio Selector: Typed Contract and Operation

PURPOSE

Model Portfolio Selector is a skills and orchestration tool whose responsibility is to rank candidates using task, evidence, cost, risk and authority signals. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-46 - Model Portfolio Selector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-46 - Model Portfolio Selector: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Model Portfolio Selector result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-46 - Model Portfolio Selector: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-46 - Model Portfolio Selector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Model Portfolio Selector: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-46 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-46 - Model Portfolio Selector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to rank candidates using task, evidence, cost, risk and authority signals. Verification design for TL-46 - Model Portfolio Selector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-46 - Model Portfolio Selector: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-46

Trigger	Name the condition that activates TL-46; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-46 - Model Portfolio Selector.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-46 - Model Portfolio Selector: Verification, Security and Recovery

VERIFICATION SUITE

Test Model Portfolio Selector with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-46 - Model Portfolio Selector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-46 - Model Portfolio Selector: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-46 - Model Portfolio Selector: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-46 - Model Portfolio Selector: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Model Portfolio Selector: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-46 - Model Portfolio Selector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-46, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-46 - Model Portfolio Selector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Model Portfolio Selector with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-46 - Model Portfolio Selector: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-46

Trigger	Name the condition that activates TL-46; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-46 - Model Portfolio Selector.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-47 - Subagent Contract Generator: Typed Contract and Operation

PURPOSE

Subagent Contract Generator is a skills and orchestration tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-47 - Subagent Contract Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-47 - Subagent Contract Generator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Subagent Contract Generator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-47 - Subagent Contract Generator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-47 - Subagent Contract Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Subagent Contract Generator: مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition. الأداة لا تملك authority

INVOCATION TRIGGER

Invoke TL-47 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-47 - Subagent Contract Generator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Lifecycle and integration for TL-47 - Subagent Contract Generator: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-47 - Subagent Contract Generator: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-47

Trigger	Name the condition that activates TL-47; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-47 - Subagent Contract Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-47 - Subagent Contract Generator: Verification, Security and Recovery

VERIFICATION SUITE

Test Subagent Contract Generator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-47 - Subagent Contract Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-47 - Subagent Contract Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-47 - Subagent Contract Generator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-47 - Subagent Contract Generator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Subagent Contract Generator: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-47 - Subagent Contract Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-47, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-47 - Subagent Contract Generator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Subagent Contract Generator with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-47 - Subagent Contract Generator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-47

Trigger	Name the condition that activates TL-47; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-47 - Subagent Contract Generator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-48 - Specialist Agent Spawner: Typed Contract and Operation

PURPOSE

Specialist Agent Spawner is a skills and orchestration tool whose responsibility is to perform bounded execution through typed permissions and observable postconditions. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-48 - Specialist Agent Spawner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-48 - Specialist Agent Spawner: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Specialist Agent Spawner result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-48 - Specialist Agent Spawner: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-48 - Specialist Agent Spawner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Specialist Agent Spawner: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-48 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-48 - Specialist Agent Spawner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to perform bounded execution through typed permissions and observable postconditions. Verification design for TL-48 - Specialist Agent Spawner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-48 - Specialist Agent Spawner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-48

Trigger	Name the condition that activates TL-48; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-48 - Specialist Agent Spawner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-48 - Specialist Agent Spawner: Verification, Security and Recovery

VERIFICATION SUITE

Test Specialist Agent Spawner with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-48 - Specialist Agent Spawner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-48 - Specialist Agent Spawner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-48 - Specialist Agent Spawner: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-48 - Specialist Agent Spawner: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Specialist Agent Spawner: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-48 - Specialist Agent Spawner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-48, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-48 - Specialist Agent Spawner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Specialist Agent Spawner with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-48 - Specialist Agent Spawner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-48

Trigger	Name the condition that activates TL-48; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-48 - Specialist Agent Spawner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-49 - Merge Arbiter: Typed Contract and Operation

PURPOSE

Merge Arbiter is a skills and orchestration tool whose responsibility is to rank candidates using task, evidence, cost, risk and authority signals. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-49 - Merge Arbiter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-49 - Merge Arbiter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Merge Arbiter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-49 - Merge Arbiter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-49 - Merge Arbiter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Merge Arbiter: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وحالته.

INVOCATION TRIGGER

Invoke TL-49 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-49 - Merge Arbiter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to rank candidates using task, evidence, cost, risk and authority signals. Failure engineering for TL-49 - Merge Arbiter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-49 - Merge Arbiter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-49

Trigger	Name the condition that activates TL-49; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-49 - Merge Arbiter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-49 - Merge Arbiter: Verification, Security and Recovery

VERIFICATION SUITE

Test Merge Arbiter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-49 - Merge Arbiter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-49 - Merge Arbiter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-49 - Merge Arbiter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-49 - Merge Arbiter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Merge Arbiter: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state وstate.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-49 - Merge Arbiter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-49, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-49 - Merge Arbiter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Merge Arbiter with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-49 - Merge Arbiter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-49

Trigger	Name the condition that activates TL-49; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-49 - Merge Arbiter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-50 - Coordination Loss Meter: Typed Contract and Operation

PURPOSE

Coordination Loss Meter is a skills and orchestration tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Security analysis for TL-50 - Coordination Loss Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Failure engineering for TL-50 - Coordination Loss Meter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED OUTPUTS

Output includes Coordination Loss Meter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Observability design for TL-50 - Coordination Loss Meter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Security analysis for TL-50 - Coordination Loss Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Coordination Loss Meter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-50 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Lifecycle and integration for TL-50 - Coordination Loss Meter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. Lifecycle and integration for TL-50 - Coordination Loss Meter: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Contract detail for TL-50 - Coordination Loss Meter: preconditions declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-50

Trigger	Name the condition that activates TL-50; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-50 - Coordination Loss Meter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-50 - Coordination Loss Meter: Verification, Security and Recovery

VERIFICATION SUITE

Test Coordination Loss Meter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Security analysis for TL-50 - Coordination Loss Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. State-transition detail for TL-50 - Coordination Loss Meter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Failure engineering for TL-50 - Coordination Loss Meter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Failure engineering for TL-50 - Coordination Loss Meter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Coordination Loss Meter: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Security analysis for TL-50 - Coordination Loss Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TELEMETRY

Record TL-50, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Security analysis for TL-50 - Coordination Loss Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

INTEGRATION TESTS

Compose Coordination Loss Meter with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. State-transition detail for TL-50 - Coordination Loss Meter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-50

Trigger	Name the condition that activates TL-50; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-50 - Coordination Loss Meter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-51 - Handoff Capsule Composer: Typed Contract and Operation

PURPOSE

Handoff Capsule Composer is a skills and orchestration tool whose responsibility is to compile structured inputs into a schema-valid artifact. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-51 - Handoff Capsule Composer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-51 - Handoff Capsule Composer: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Handoff Capsule Composer result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-51 - Handoff Capsule Composer: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-51 - Handoff Capsule Composer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Handoff Capsule Composer: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-51 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-51 - Handoff Capsule Composer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to compile structured inputs into a schema-valid artifact. Observability design for TL-51 - Handoff Capsule Composer: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-51 - Handoff Capsule Composer: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX — TL-51

Trigger	Name the condition that activates TL-51; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-51 - Handoff Capsule Composer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-51 - Handoff Capsule Composer: Verification, Security and Recovery

VERIFICATION SUITE

Test Handoff Capsule Composer with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-51 - Handoff Capsule Composer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-51 - Handoff Capsule Composer: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-51 - Handoff Capsule Composer: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-51 - Handoff Capsule Composer: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Handoff Capsule Composer: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-51 - Handoff Capsule Composer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-51, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-51 - Handoff Capsule Composer: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Handoff Capsule Composer with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-51 - Handoff Capsule Composer: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-51

Trigger	Name the condition that activates TL-51; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-51 - Handoff Capsule Composer.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-52 - Bounded Loop Controller: Typed Contract and Operation

PURPOSE

Bounded Loop Controller is a skills and orchestration tool whose responsibility is to perform bounded execution through typed permissions and observable postconditions. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-52 - Bounded Loop Controller: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-52 - Bounded Loop Controller: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes Bounded Loop Controller result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-52 - Bounded Loop Controller: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MHSG, MSHR, MHSO, MBHL. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-52 - Bounded Loop Controller: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Bounded Loop Controller: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-52 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-52 - Bounded Loop Controller: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to perform bounded execution through typed permissions and observable postconditions. Contract detail for TL-52 - Bounded Loop Controller: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-52 - Bounded Loop Controller: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX

TL-52

Trigger	Name the condition that activates TL-52; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-52 - Bounded Loop Controller.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-52 - Bounded Loop Controller: Verification, Security and Recovery

VERIFICATION SUITE

Test Bounded Loop Controller with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-52 - Bounded Loop Controller: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-52 - Bounded Loop Controller: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-52 - Bounded Loop Controller: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-52 - Bounded Loop Controller: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Bounded Loop Controller: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-52 - Bounded Loop Controller: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-52, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-52 - Bounded Loop Controller: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose Bounded Loop Controller with MHSG, MSHR, MHSO, MBHL, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-52 - Bounded Loop Controller: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-52

Trigger	Name the condition that activates TL-52; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-52 - Bounded Loop Controller.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-53 - Tool Affordance Registry: Typed Contract and Operation

PURPOSE

Tool Affordance Registry is a tool safety and courts tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-53 - Tool Affordance Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-53 - Tool Affordance Registry: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes Tool Affordance Registry result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-53 - Tool Affordance Registry: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-53 - Tool Affordance Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Tool Affordance Registry: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-53 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-53 - Tool Affordance Registry: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. Contract detail for TL-53 - Tool Affordance Registry: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-53 - Tool Affordance Registry: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX — TL-53

Trigger	Name the condition that activates TL-53; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-53 - Tool Affordance Registry.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-53 - Tool Affordance Registry: Verification, Security and Recovery

VERIFICATION SUITE

Test Tool Affordance Registry with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-53 - Tool Affordance Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-53 - Tool Affordance Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-53 - Tool Affordance Registry: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-53 - Tool Affordance Registry: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Tool Affordance Registry: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-53 - Tool Affordance Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-53, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-53 - Tool Affordance Registry: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose Tool Affordance Registry with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-53 - Tool Affordance Registry: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-53

Trigger	Name the condition that activates TL-53; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-53 - Tool Affordance Registry.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-54 - Tool Sandbox Executor: Typed Contract and Operation

PURPOSE

Tool Sandbox Executor is a tool safety and courts tool whose responsibility is to perform bounded execution through typed permissions and observable postconditions. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-54 - Tool Sandbox Executor: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-54 - Tool Sandbox Executor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Tool Sandbox Executor result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-54 - Tool Sandbox Executor: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-54 - Tool Sandbox Executor: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Tool Sandbox Executor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-54 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-54 - Tool Sandbox Executor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to perform bounded execution through typed permissions and observable postconditions. Security analysis for TL-54 - Tool Sandbox Executor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-54 - Tool Sandbox Executor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-54

Trigger	Name the condition that activates TL-54; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-54 - Tool Sandbox Executor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-54 - Tool Sandbox Executor: Verification, Security and Recovery

VERIFICATION SUITE

Test Tool Sandbox Executor with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-54 - Tool Sandbox Executor: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-54 - Tool Sandbox Executor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-54 - Tool Sandbox Executor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-54 - Tool Sandbox Executor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Tool Sandbox Executor: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-54 - Tool Sandbox Executor: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-54, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-54 - Tool Sandbox Executor: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Tool Sandbox Executor with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-54 - Tool Sandbox Executor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-54

Trigger	Name the condition that activates TL-54; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-54 - Tool Sandbox Executor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-55 - Injection Sentinel: Typed Contract and Operation

PURPOSE

Injection Sentinel is a tool safety and courts tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-55 - Injection Sentinel: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-55 - Injection Sentinel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Injection Sentinel result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-55 - Injection Sentinel: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-55 - Injection Sentinel: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Injection Sentinel: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-55 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-55 - Injection Sentinel: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Failure engineering for TL-55 - Injection Sentinel: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-55 - Injection Sentinel: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-55

Trigger	Name the condition that activates TL-55; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-55 - Injection Sentinel.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-55 - Injection Sentinel: Verification, Security and Recovery

VERIFICATION SUITE

Test Injection Sentinel with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-55 - Injection Sentinel: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-55 - Injection Sentinel: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-55 - Injection Sentinel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-55 - Injection Sentinel: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Injection Sentinel: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-55 - Injection Sentinel: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-55, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-55 - Injection Sentinel: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Injection Sentinel with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-55 - Injection Sentinel: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-55

Trigger	Name the condition that activates TL-55; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-55 - Injection Sentinel.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-56 - Permission Scope Checker: Typed Contract and Operation

PURPOSE

Permission Scope Checker is a tool safety and courts tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-56 - Permission Scope Checker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-56 - Permission Scope Checker: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Permission Scope Checker result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-56 - Permission Scope Checker: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-56 - Permission Scope Checker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Permission Scope Checker: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-56 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-56 - Permission Scope Checker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Failure engineering for TL-56 - Permission Scope Checker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-56 - Permission Scope Checker: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX — TL-56

Trigger	Name the condition that activates TL-56; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-56 - Permission Scope Checker.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-56 - Permission Scope Checker: Verification, Security and Recovery

VERIFICATION SUITE

Test Permission Scope Checker with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-56 - Permission Scope Checker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-56 - Permission Scope Checker: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-56 - Permission Scope Checker: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-56 - Permission Scope Checker: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Permission Scope Checker: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-56 - Permission Scope Checker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-56, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-56 - Permission Scope Checker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Permission Scope Checker with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-56 - Permission Scope Checker: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-56

Trigger	Name the condition that activates TL-56; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-56 - Permission Scope Checker.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-57 - Precondition Validator: Typed Contract and Operation

PURPOSE

Precondition Validator is a tool safety and courts tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-57 - Precondition Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-57 - Precondition Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Precondition Validator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-57 - Precondition Validator: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-57 - Precondition Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Precondition Validator: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ postcondition وreturned state.

INVOCATION TRIGGER

Invoke TL-57 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-57 - Precondition Validator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. State-transition detail for TL-57 - Precondition Validator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-57 - Precondition Validator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-57

Trigger	Name the condition that activates TL-57; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-57 - Precondition Validator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-57 - Precondition Validator: Verification, Security and Recovery

VERIFICATION SUITE

Test Precondition Validator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-57 - Precondition Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-57 - Precondition Validator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-57 - Precondition Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-57 - Precondition Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Precondition Validator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-57 - Precondition Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-57, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-57 - Precondition Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Precondition Validator with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-57 - Precondition Validator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-57

Trigger	Name the condition that activates TL-57; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-57 - Precondition Validator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-58 - Output Schema Validator: Typed Contract and Operation

PURPOSE

Output Schema Validator is a tool safety and courts tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-58 - Output Schema Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-58 - Output Schema Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Output Schema Validator result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-58 - Output Schema Validator: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-58 - Output Schema Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Output Schema Validator: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-58 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-58 - Output Schema Validator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. State-transition detail for TL-58 - Output Schema Validator: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-58 - Output Schema Validator: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-58

Trigger	Name the condition that activates TL-58; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-58 - Output Schema Validator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-58 - Output Schema Validator: Verification, Security and Recovery

VERIFICATION SUITE

Test Output Schema Validator with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-58 - Output Schema Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-58 - Output Schema Validator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-58 - Output Schema Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-58 - Output Schema Validator: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Output Schema Validator: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-58 - Output Schema Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-58, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-58 - Output Schema Validator: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Output Schema Validator with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-58 - Output Schema Validator: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-58

Trigger	Name the condition that activates TL-58; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-58 - Output Schema Validator.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-59 - Falsification Court Runner: Typed Contract and Operation

PURPOSE

Falsification Court Runner is a tool safety and courts tool whose responsibility is to perform bounded execution through typed permissions and observable postconditions. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-59 - Falsification Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-59 - Falsification Court Runner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Falsification Court Runner result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-59 - Falsification Court Runner: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-59 - Falsification Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Falsification Court Runner: الأداة لا تملك authority بعد فقط يثبت نجاحها؛ مستقلة returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-59 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-59 - Falsification Court Runner: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to perform bounded execution through typed permissions and observable postconditions. Failure engineering for TL-59 - Falsification Court Runner: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-59 - Falsification Court Runner: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-59

Trigger	Name the condition that activates TL-59; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-59 - Falsification Court Runner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-59 - Falsification Court Runner: Verification, Security and Recovery

VERIFICATION SUITE

Test Falsification Court Runner with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-59 - Falsification Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-59 - Falsification Court Runner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-59 - Falsification Court Runner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-59 - Falsification Court Runner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Falsification Court Runner: مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-59 - Falsification Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-59, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-59 - Falsification Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Falsification Court Runner with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-59 - Falsification Court Runner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-59

Trigger	Name the condition that activates TL-59; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-59 - Falsification Court Runner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-60 - Gap-Filling Court Runner: Typed Contract and Operation

PURPOSE

Gap-Filling Court Runner is a tool safety and courts tool whose responsibility is to perform bounded execution through typed permissions and observable postconditions. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-60 - Gap-Filling Court Runner: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-60 - Gap-Filling Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Gap-Filling Court Runner result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-60 - Gap-Filling Court Runner: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MTIT, MIRTT, MCIT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-60 - Gap-Filling Court Runner: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Gap-Filling Court Runner: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-60 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-60 - Gap-Filling Court Runner: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to perform bounded execution through typed permissions and observable postconditions. Security analysis for TL-60 - Gap-Filling Court Runner: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-60 - Gap-Filling Court Runner: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-60

Trigger	Name the condition that activates TL-60; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-60 - Gap-Filling Court Runner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-60 - Gap-Filling Court Runner: Verification, Security and Recovery

VERIFICATION SUITE

Test Gap-Filling Court Runner with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-60 - Gap-Filling Court Runner: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-60 - Gap-Filling Court Runner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-60 - Gap-Filling Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-60 - Gap-Filling Court Runner: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Gap-Filling Court Runner: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-60 - Gap-Filling Court Runner: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-60, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-60 - Gap-Filling Court Runner: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Gap-Filling Court Runner with MTIT, MIRTT, MCIT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-60 - Gap-Filling Court Runner: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-60

Trigger	Name the condition that activates TL-60; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-60 - Gap-Filling Court Runner.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-61 - Benchmark Tribunal: Typed Contract and Operation

PURPOSE

Benchmark Tribunal is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-61 - Benchmark Tribunal: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-61 - Benchmark Tribunal: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Benchmark Tribunal result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-61 - Benchmark Tribunal: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-61 - Benchmark Tribunal: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Benchmark Tribunal: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-61 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-61 - Benchmark Tribunal: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Failure engineering for TL-61 - Benchmark Tribunal: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-61 - Benchmark Tribunal: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-61

Trigger	Name the condition that activates TL-61; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-61 - Benchmark Tribunal.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-61 - Benchmark Tribunal: Verification, Security and Recovery

VERIFICATION SUITE

Test Benchmark Tribunal with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-61 - Benchmark Tribunal: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-61 - Benchmark Tribunal: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-61 - Benchmark Tribunal: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-61 - Benchmark Tribunal: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Benchmark Tribunal: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-61 - Benchmark Tribunal: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-61, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-61 - Benchmark Tribunal: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Benchmark Tribunal with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-61 - Benchmark Tribunal: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-61

Trigger	Name the condition that activates TL-61; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-61 - Benchmark Tribunal.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-62 - Ablation Harness: Typed Contract and Operation

PURPOSE

Ablation Harness is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-62 - Ablation Harness: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-62 - Ablation Harness: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes Ablation Harness result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-62 - Ablation Harness: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-62 - Ablation Harness: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Ablation Harness: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-62 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-62 - Ablation Harness: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Contract detail for TL-62 - Ablation Harness: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-62 - Ablation Harness: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX — TL-62

Trigger	Name the condition that activates TL-62; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-62 - Ablation Harness.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-62 - Ablation Harness: Verification, Security and Recovery

VERIFICATION SUITE

Test Ablation Harness with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-62 - Ablation Harness: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-62 - Ablation Harness: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-62 - Ablation Harness: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-62 - Ablation Harness: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Ablation Harness: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-62 - Ablation Harness: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-62, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-62 - Ablation Harness: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose Ablation Harness with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-62 - Ablation Harness: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-62

Trigger	Name the condition that activates TL-62; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-62 - Ablation Harness.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-63 - Trajectory Recorder: Typed Contract and Operation

PURPOSE

Trajectory Recorder is a benchmark, governance and release tool whose responsibility is to persist a versioned record with provenance, scope and retention policy. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-63 - Trajectory Recorder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-63 - Trajectory Recorder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Trajectory Recorder result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-63 - Trajectory Recorder: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-63 - Trajectory Recorder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Trajectory Recorder: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-63 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-63 - Trajectory Recorder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to persist a versioned record with provenance, scope and retention policy. Failure engineering for TL-63 - Trajectory Recorder: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-63 - Trajectory Recorder: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-63

Trigger	Name the condition that activates TL-63; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-63 - Trajectory Recorder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-63 - Trajectory Recorder: Verification, Security and Recovery

VERIFICATION SUITE

Test Trajectory Recorder with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-63 - Trajectory Recorder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-63 - Trajectory Recorder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-63 - Trajectory Recorder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-63 - Trajectory Recorder: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Trajectory Recorder: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-63 - Trajectory Recorder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-63, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-63 - Trajectory Recorder: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Trajectory Recorder with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-63 - Trajectory Recorder: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-63

Trigger	Name the condition that activates TL-63; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-63 - Trajectory Recorder.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-64 - Cost and Latency Meter: Typed Contract and Operation

PURPOSE

Cost and Latency Meter is a benchmark, governance and release tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-64 - Cost and Latency Meter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-64 - Cost and Latency Meter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Cost and Latency Meter result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-64 - Cost and Latency Meter: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-64 - Cost and Latency Meter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Cost and Latency Meter: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-64 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-64 - Cost and Latency Meter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. State-transition detail for TL-64 - Cost and Latency Meter: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-64 - Cost and Latency Meter: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-64

Trigger	Name the condition that activates TL-64; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-64 - Cost and Latency Meter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-64 - Cost and Latency Meter: Verification, Security and Recovery

VERIFICATION SUITE

Test Cost and Latency Meter with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-64 - Cost and Latency Meter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-64 - Cost and Latency Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-64 - Cost and Latency Meter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-64 - Cost and Latency Meter: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Cost and Latency Meter: الأداة لا تملك authority بعد فقط يثبت نجاحها مستقلة؛ postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-64 - Cost and Latency Meter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-64, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-64 - Cost and Latency Meter: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Cost and Latency Meter with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-64 - Cost and Latency Meter: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-64

Trigger	Name the condition that activates TL-64; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-64 - Cost and Latency Meter.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-65 - Risk Tier Classifier: Typed Contract and Operation

PURPOSE

Risk Tier Classifier is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-65 - Risk Tier Classifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-65 - Risk Tier Classifier: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes Risk Tier Classifier result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-65 - Risk Tier Classifier: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-65 - Risk Tier Classifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Risk Tier Classifier: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-65 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-65 - Risk Tier Classifier: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Contract detail for TL-65 - Risk Tier Classifier: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-65 - Risk Tier Classifier: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX — TL-65

Trigger	Name the condition that activates TL-65; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-65 - Risk Tier Classifier.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-65 - Risk Tier Classifier: Verification, Security and Recovery

VERIFICATION SUITE

Test Risk Tier Classifier with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-65 - Risk Tier Classifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-65 - Risk Tier Classifier: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-65 - Risk Tier Classifier: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-65 - Risk Tier Classifier: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Risk Tier Classifier: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-65 - Risk Tier Classifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-65, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-65 - Risk Tier Classifier: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose Risk Tier Classifier with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-65 - Risk Tier Classifier: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-65

Trigger	Name the condition that activates TL-65; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-65 - Risk Tier Classifier.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-66 - Human Approval Gate: Typed Contract and Operation

PURPOSE

Human Approval Gate is a benchmark, governance and release tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-66 - Human Approval Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-66 - Human Approval Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Human Approval Gate result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-66 - Human Approval Gate: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-66 - Human Approval Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Human Approval Gate: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-66 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-66 - Human Approval Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Observability design for TL-66 - Human Approval Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-66 - Human Approval Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX — TL-66

Trigger	Name the condition that activates TL-66; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-66 - Human Approval Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-66 - Human Approval Gate: Verification, Security and Recovery

VERIFICATION SUITE

Test Human Approval Gate with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-66 - Human Approval Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-66 - Human Approval Gate: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-66 - Human Approval Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-66 - Human Approval Gate: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Human Approval Gate: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-66 - Human Approval Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-66, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-66 - Human Approval Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Human Approval Gate with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-66 - Human Approval Gate: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-66

Trigger	Name the condition that activates TL-66; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-66 - Human Approval Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-67 - Rollback Manager: Typed Contract and Operation

PURPOSE

Rollback Manager is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Observability design for TL-67 - Rollback Manager: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Verification design for TL-67 - Rollback Manager: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED OUTPUTS

Output includes Rollback Manager result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Contract detail for TL-67 - Rollback Manager: typed outputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Observability design for TL-67 - Rollback Manager: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Rollback Manager: الأداة لا تملك authority بعد مستقلة؛ نجاحها يثبت فقط returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-67 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. State-transition detail for TL-67 - Rollback Manager: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. State-transition detail for TL-67 - Rollback Manager: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Failure engineering for TL-67 - Rollback Manager: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-67

Trigger	Name the condition that activates TL-67; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-67 - Rollback Manager.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-67 - Rollback Manager: Verification, Security and Recovery

VERIFICATION SUITE

Test Rollback Manager with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Observability design for TL-67 - Rollback Manager: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Security analysis for TL-67 - Rollback Manager: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Verification design for TL-67 - Rollback Manager: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Verification design for TL-67 - Rollback Manager: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Rollback Manager: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Observability design for TL-67 - Rollback Manager: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TELEMETRY

Record TL-67, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Observability design for TL-67 - Rollback Manager: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

INTEGRATION TESTS

Compose Rollback Manager with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Security analysis for TL-67 - Rollback Manager: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-67

Trigger	Name the condition that activates TL-67; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-67 - Rollback Manager.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-68 - Secret Redactor: Typed Contract and Operation

PURPOSE

Secret Redactor is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. State-transition detail for TL-68 - Secret Redactor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Contract detail for TL-68 - Secret Redactor: typed inputs declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED OUTPUTS

Output includes Secret Redactor result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Security analysis for TL-68 - Secret Redactor: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. State-transition detail for TL-68 - Secret Redactor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

Secret Redactor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-68 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Verification design for TL-68 - Secret Redactor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Verification design for TL-68 - Secret Redactor: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Observability design for TL-68 - Secret Redactor: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-68

Trigger	Name the condition that activates TL-68; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-68 - Secret Redactor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-68 - Secret Redactor: Verification, Security and Recovery

VERIFICATION SUITE

Test Secret Redactor with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. State-transition detail for TL-68 - Secret Redactor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Lifecycle and integration for TL-68 - Secret Redactor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Contract detail for TL-68 - Secret Redactor: rollback and recovery declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Contract detail for TL-68 - Secret Redactor: promotion gate declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Secret Redactor: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. State-transition detail for TL-68 - Secret Redactor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TELEMETRY

Record TL-68, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. State-transition detail for TL-68 - Secret Redactor: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

INTEGRATION TESTS

Compose Secret Redactor with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Lifecycle and integration for TL-68 - Secret Redactor: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-68

Trigger	Name the condition that activates TL-68; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-68 - Secret Redactor.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-69 - Data Lineage Tracker: Typed Contract and Operation

PURPOSE

Data Lineage Tracker is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Failure engineering for TL-69 - Data Lineage Tracker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. State-transition detail for TL-69 - Data Lineage Tracker: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

TYPED OUTPUTS

Output includes Data Lineage Tracker result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Verification design for TL-69 - Data Lineage Tracker: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Failure engineering for TL-69 - Data Lineage Tracker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

Data Lineage Tracker: الأداة لا تملك authority بعد فقط يثبت نجاحها؛ مستقلة returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-69 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Observability design for TL-69 - Data Lineage Tracker: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Observability design for TL-69 - Data Lineage Tracker: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Lifecycle and integration for TL-69 - Data Lineage Tracker: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

OPERATIONAL EVIDENCE MATRIX

TL-69

Trigger	Name the condition that activates TL-69; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-69 - Data Lineage Tracker.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-69 - Data Lineage Tracker: Verification, Security and Recovery

VERIFICATION SUITE

Test Data Lineage Tracker with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Failure engineering for TL-69 - Data Lineage Tracker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Contract detail for TL-69 - Data Lineage Tracker: security controls declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. State-transition detail for TL-69 - Data Lineage Tracker: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. State-transition detail for TL-69 - Data Lineage Tracker: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك :Data Lineage Tracker
postcondition وreturned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Failure engineering for TL-69 - Data Lineage Tracker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

TELEMETRY

Record TL-69, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Failure engineering for TL-69 - Data Lineage Tracker: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

INTEGRATION TESTS

Compose Data Lineage Tracker with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Contract detail for TL-69 - Data Lineage Tracker: integration tests declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

OPERATIONAL EVIDENCE MATRIX

TL-69

Trigger	Name the condition that activates TL-69; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-69 - Data Lineage Tracker.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-70 - Observability Dashboard: Typed Contract and Operation

PURPOSE

Observability Dashboard is a benchmark, governance and release tool whose responsibility is to transform a governed task state into an auditable operational object. It is not an autonomous authority; it operates only inside the active task contract. Lifecycle and integration for TL-70 - Observability Dashboard: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Observability design for TL-70 - Observability Dashboard: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

TYPED OUTPUTS

Output includes Observability Dashboard result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. State-transition detail for TL-70 - Observability Dashboard: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Lifecycle and integration for TL-70 - Observability Dashboard: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Observability Dashboard: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-70 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Failure engineering for TL-70 - Observability Dashboard: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to transform a governed task state into an auditable operational object. Failure engineering for TL-70 - Observability Dashboard: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Security analysis for TL-70 - Observability Dashboard: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

OPERATIONAL EVIDENCE MATRIX

TL-70

Trigger	Name the condition that activates TL-70; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-70 - Observability Dashboard.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-70 - Observability Dashboard: Verification, Security and Recovery

VERIFICATION SUITE

Test Observability Dashboard with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Lifecycle and integration for TL-70 - Observability Dashboard: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Verification design for TL-70 - Observability Dashboard: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Observability design for TL-70 - Observability Dashboard: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Observability design for TL-70 - Observability Dashboard: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

Observability Dashboard: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Lifecycle and integration for TL-70 - Observability Dashboard: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TELEMETRY

Record TL-70, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Lifecycle and integration for TL-70 - Observability Dashboard: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

INTEGRATION TESTS

Compose Observability Dashboard with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Verification design for TL-70 - Observability Dashboard: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-70

Trigger	Name the condition that activates TL-70; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-70 - Observability Dashboard.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-71 - Drift Detector: Typed Contract and Operation

PURPOSE

Drift Detector is a benchmark, governance and release tool whose responsibility is to measure state, quality, risk or drift and return calibrated diagnostics. It is not an autonomous authority; it operates only inside the active task contract. Verification design for TL-71 - Drift Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Security analysis for TL-71 - Drift Detector: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

TYPED OUTPUTS

Output includes Drift Detector result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Lifecycle and integration for TL-71 - Drift Detector: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Verification design for TL-71 - Drift Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

Drift Detector: postcondition مستقلة؛ نجاحها يثبت فقط بعد authority الأداة لا تملك وreturned state.

INVOCATION TRIGGER

Invoke TL-71 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Contract detail for TL-71 - Drift Detector: invocation trigger declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to measure state, quality, risk or drift and return calibrated diagnostics. Contract detail for TL-71 - Drift Detector: algorithm declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. State-transition detail for TL-71 - Drift Detector: capture pre-state, calculate the intended delta, execute once under bounded authority, then read authoritative returned state. Partial success is represented separately from success and failure. Concurrent or duplicate calls use idempotency and conflict policy. A postcondition mismatch blocks downstream actions. The tool never infers that an email, deployment, file write or transaction succeeded merely because a request was issued; it requires the external system response or a later reconciliation probe.

OPERATIONAL EVIDENCE MATRIX — TL-71

Trigger	Name the condition that activates TL-71; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-71 - Drift Detector.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-71 - Drift Detector: Verification, Security and Recovery

VERIFICATION SUITE

Test Drift Detector with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Verification design for TL-71 - Drift Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Failure engineering for TL-71 - Drift Detector: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Security analysis for TL-71 - Drift Detector: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Security analysis for TL-71 - Drift Detector: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

Drift Detector: الأداة لا تملك authority بعد postcondition مستقلة؛ نجاحها يثبت فقط returned state.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Verification design for TL-71 - Drift Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

TELEMETRY

Record TL-71, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Verification design for TL-71 - Drift Detector: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

INTEGRATION TESTS

Compose Drift Detector with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Failure engineering for TL-71 - Drift Detector: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

OPERATIONAL EVIDENCE MATRIX

TL-71

Trigger	Name the condition that activates TL-71; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-71 - Drift Detector.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-72 - Release Promotion Gate: Typed Contract and Operation

PURPOSE

Release Promotion Gate is a benchmark, governance and release tool whose responsibility is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. It is not an autonomous authority; it operates only inside the active task contract. Contract detail for TL-72 - Release Promotion Gate: purpose declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TYPED INPUTS

Inputs include object identifiers, schema version, task and goal references, evidence pointers, world-state timestamp, authority scope, risk tier, resource budget and idempotency key. Lifecycle and integration for TL-72 - Release Promotion Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

TYPED OUTPUTS

Output includes Release Promotion Gate result, status, evidence pointers, confidence or quality score, mutation record, elapsed cost, warnings, postcondition result and rollback state. Failure engineering for TL-72 - Release Promotion Gate: enumerate invalid input, missing dependency, stale read, timeout, rate limit, network partition, partial mutation, duplicate execution, permission loss, corrupted output and unavailable rollback. Each signature maps to STOP, RETRY-WITH-CHANGE, DEGRADE, HUMAN-REVIEW or ROLLBACK. Retry counts and backoff are bounded to prevent amplification. The incident record preserves enough state for diagnosis without copying secrets or sensitive payloads into logs.

THEORY MAPPING

Primary theoretical dependencies: MDBCT, MAOLT, MRRMT. The tool is evidence that a theory is operational only when behavior and tests differ from the baseline. Contract detail for TL-72 - Release Promotion Gate: theory mapping declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

Release Promotion Gate: الأداة لا تملك authority بعد نجاحها يثبت فقط مستقلة؛ returned state وpostcondition.

INVOCATION TRIGGER

Invoke TL-72 when the required object is missing, stale, invalid or decision-critical. Skip it when a cached validated result remains applicable or when the task lacks authority. Security analysis for TL-72 - Release Promotion Gate: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

ALGORITHM

Validate preconditions; normalize inputs; execute the minimum bounded transformation; collect returned state; validate schema and postconditions; emit telemetry; then commit or rollback. The core action is to evaluate an artifact against explicit constraints and emit a pass, hold or fail verdict. Security analysis for TL-72 - Release Promotion Gate: separate instruction from untrusted data, propagate taint labels and minimize tool permissions. Resolve secrets through references; do not embed them in prompts, artifacts or telemetry. Validate targets immediately before consequential action. Test indirect injection, path manipulation, malicious documents, confused-deputy behavior and permission revocation. If authorization is ambiguous, the correct output is a blocked action plus a precise request for human authority.

PRECONDITIONS

Required schema is known; authority is valid; dependencies are reachable; data policy permits processing; timeout and stop conditions exist; expected state is captured before mutation. Verification design for TL-72 - Release Promotion Gate: run schema, property, mutation, integration, concurrency, timeout, adversarial and rollback tests. The oracle checks both returned object and world state. Use fixtures for reproducibility and live probes only when authorized. Compare the tool against a manual or simpler baseline and measure defect discovery, false pass, false block, latency, cost and recovery. A tool is production-ready only when the failure path is as specified as the success path.

OPERATIONAL EVIDENCE MATRIX

TL-72

Trigger	Name the condition that activates TL-72; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-72 - Release Promotion Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

TL-72 - Release Promotion Gate: Verification, Security and Recovery

VERIFICATION SUITE

Test Release Promotion Gate with clean, malformed, empty, stale, conflicting, adversarial and high-volume inputs. Verify deterministic schema behavior where required and calibrated uncertainty elsewhere. Contract detail for TL-72 - Release Promotion Gate: verification suite declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

SECURITY CONTROLS

Use least privilege, sandboxing, secret references, taint tracking, data minimization, allowlists and action-time permission checks. Untrusted retrieved text never becomes privileged instruction. Observability design for TL-72 - Release Promotion Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

ROLLBACK AND RECOVERY

Prefer idempotent operations. Capture pre-state and mutation identifier. On failed postcondition, stop dependent actions, revert if safe, record residual change and escalate to a human when reversal is uncertain. Lifecycle and integration for TL-72 - Release Promotion Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

PROMOTION GATE

Production-ready requires typed I/O, permission tests, timeout behavior, retry limits, postcondition validation, telemetry, replay, adversarial coverage, rollback rehearsal and an accountable owner. Lifecycle and integration for TL-72 - Release Promotion Gate: document owner, dependency policy, release channel, canary cohort, compatibility range and deprecation path. Integration tests verify the upstream contract and every downstream consumer. Promotion requires a replayable test pack and rollback rehearsal. Drift detection compares input distribution, output quality, cost and failure signatures across versions. A newer tool is not preferred unless its measured value exceeds the migration and governance cost.

Release Promotion Gate: الأداة لا تملك authority مستقلة؛ نجاحها يثبت فقط بعد returned state وpostcondition.

FAILURE SIGNATURES

Detect invalid input, missing authority, dependency failure, timeout, partial mutation, stale read, duplicate execution, inconsistent returned state, privacy leakage and false-success reporting. Contract detail for TL-72 - Release Promotion Gate: failure signatures declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

TELEMETRY

Record TL-72, tool and schema version, trigger, sanitized inputs, route, latency, tokens or compute, returned status, postcondition, court verdict and rollback. Do not expose secrets. Contract detail for TL-72 - Release Promotion Gate: telemetry declares schema version, required and optional fields, authority scope, idempotency key, risk tier, budget, timeout and expected postcondition. Unknown fields and lossy coercions are handled explicitly. The caller receives machine-readable status, warnings, evidence pointers and residual uncertainty. Validation failure occurs before external mutation. Compatibility tests cover previous schema versions and migration behavior so a release cannot silently alter the meaning of stored records.

INTEGRATION TESTS

Compose Release Promotion Gate with MDBCT, MAOLT, MRRMT, then remove it in ablation. Measure quality, cost, latency, authority violations, recovery time and downstream defect rate. Observability design for TL-72 - Release Promotion Gate: log tool and schema version, trigger, sanitized inputs, route, start and end timestamps, dependency versions, cost, returned status, postcondition, court verdict, mutation ID and rollback result. Link records through stable identifiers. High-cardinality or sensitive values are referenced rather than copied. Define SLOs for success, latency, error, recovery and freshness; alerting is based on actionable thresholds, not raw event volume.

OPERATIONAL EVIDENCE MATRIX

TL-72

Trigger	Name the condition that activates TL-72; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by TL-72 - Release Promotion Gate.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-01 - Cognitive Architecture and Problem Framing: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-01 develops transferable capability in Cognitive Architecture and Problem Framing. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Architecture analysis for D-01 - Cognitive Architecture and Problem Framing: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Measurement design for D-01 - Cognitive Architecture and Problem Framing: assess capability contract through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Transfer analysis for D-01 - Cognitive Architecture and Problem Framing: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Scenario design for D-01 - Cognitive Architecture and Problem Framing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for human role. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Target design for D-01 - Cognitive Architecture and Problem Framing: twenty target pattern is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Measurement design for D-01 - Cognitive Architecture and Problem Framing: assess knowledge and tools through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

TRANSFER REQUIREMENT

A capability in Cognitive Architecture and Problem Framing is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Measurement design for D-01 - Cognitive Architecture and Problem Framing: assess transfer requirement through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-01

Trigger	Name the condition that activates D-01; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-01 - Cognitive Architecture and Problem Framing.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-01 - Cognitive Architecture and Problem Framing: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Cognitive Architecture and Problem Framing. Hide some cases to reduce prompt overfitting. Architecture analysis for D-01 - Cognitive Architecture and Problem Framing: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Measurement design for D-01 - Cognitive Architecture and Problem Framing: assess metrics through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Scenario design for D-01 - Cognitive Architecture and Problem Framing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for adversarial coverage. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Cognitive Architecture and Problem Framing, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Measurement design for D-01 - Cognitive Architecture and Problem Framing: assess open research through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Scenario design for D-01 - Cognitive Architecture and Problem Framing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for baselines. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Measurement design for D-01 - Cognitive Architecture and Problem Framing: assess ablations through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Governance analysis for D-01 - Cognitive Architecture and Problem Framing: connect promotion decision to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-01

Trigger	Name the condition that activates D-01; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-01 - Cognitive Architecture and Problem Framing.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-02 - Q-MCD Question-Governed Reasoning: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-02 develops transferable capability in Q-MCD Question-Governed Reasoning. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-02 - Q-MCD Question-Governed Reasoning: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-02 - Q-MCD Question-Governed Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-02 - Q-MCD Question-Governed Reasoning: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-02 - Q-MCD Question-Governed Reasoning: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-02 - Q-MCD Question-Governed Reasoning: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-02 - Q-MCD Question-Governed Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in Q-MCD Question-Governed Reasoning is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-02 - Q-MCD Question-Governed Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-02

Trigger	Name the condition that activates D-02; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-02 - Q-MCD Question-Governed Reasoning.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-02 - Q-MCD Question-Governed Reasoning: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Q-MCD Question-Governed Reasoning. Hide some cases to reduce prompt overfitting. Measurement design for D-02 - Q-MCD Question-Governed Reasoning: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-02 - Q-MCD Question-Governed Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-02 - Q-MCD Question-Governed Reasoning: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Q-MCD Question-Governed Reasoning, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-02 - Q-MCD Question-Governed Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-02 - Q-MCD Question-Governed Reasoning: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-02 - Q-MCD Question-Governed Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-02 - Q-MCD Question-Governed Reasoning: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-02

Trigger	Name the condition that activates D-02; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-02 - Q-MCD Question-Governed Reasoning.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-03 - Long-Horizon Planning and Bounded Hyperloops: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-03 develops transferable capability in Long-Horizon Planning and Bounded Hyperloops. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-03 - Long-Horizon Planning and Bounded Hyperloops: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-03 - Long-Horizon Planning and Bounded Hyperloops: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-03 - Long-Horizon Planning and Bounded Hyperloops: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in Long-Horizon Planning and Bounded Hyperloops is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-03

Trigger	Name the condition that activates D-03; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-03 - Long-Horizon Planning and Bounded Hyperloops.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-03 - Long-Horizon Planning and Bounded Hyperloops: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Long-Horizon Planning and Bounded Hyperloops. Hide some cases to reduce prompt overfitting. Target design for D-03 - Long-Horizon Planning and Bounded Hyperloops: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-03 - Long-Horizon Planning and Bounded Hyperloops: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Long-Horizon Planning and Bounded Hyperloops, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-03 - Long-Horizon Planning and Bounded Hyperloops: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-03 - Long-Horizon Planning and Bounded Hyperloops: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-03

Trigger	Name the condition that activates D-03; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-03 - Long-Horizon Planning and Bounded Hyperloops.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-04 - Sparse HyperRouting and Model Portfolio Orchestration: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-04 develops transferable capability in Sparse HyperRouting and Model Portfolio Orchestration. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in Sparse HyperRouting and Model Portfolio Orchestration is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-04

Trigger	Name the condition that activates D-04; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-04 - Sparse HyperRouting and Model Portfolio Orchestration.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-04 - Sparse HyperRouting and Model Portfolio Orchestration: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Sparse HyperRouting and Model Portfolio Orchestration. Hide some cases to reduce prompt overfitting. Target design for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Sparse HyperRouting and Model Portfolio Orchestration, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-04 - Sparse HyperRouting and Model Portfolio Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-04

Trigger	Name the condition that activates D-04; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-04 - Sparse HyperRouting and Model Portfolio Orchestration.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-05 - Multi-Agent and Subagent Orchestration: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-05 develops transferable capability in Multi-Agent and Subagent Orchestration. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-05 - Multi-Agent and Subagent Orchestration: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-05 - Multi-Agent and Subagent Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-05 - Multi-Agent and Subagent Orchestration: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-05 - Multi-Agent and Subagent Orchestration: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-05 - Multi-Agent and Subagent Orchestration: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-05 - Multi-Agent and Subagent Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in Multi-Agent and Subagent Orchestration is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-05 - Multi-Agent and Subagent Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-05

Trigger	Name the condition that activates D-05; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-05 - Multi-Agent and Subagent Orchestration.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-05 - Multi-Agent and Subagent Orchestration: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Multi-Agent and Subagent Orchestration. Hide some cases to reduce prompt overfitting. Measurement design for D-05 - Multi-Agent and Subagent Orchestration: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-05 - Multi-Agent and Subagent Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-05 - Multi-Agent and Subagent Orchestration: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Multi-Agent and Subagent Orchestration, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-05 - Multi-Agent and Subagent Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-05 - Multi-Agent and Subagent Orchestration: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-05 - Multi-Agent and Subagent Orchestration: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-05 - Multi-Agent and Subagent Orchestration: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-05

Trigger	Name the condition that activates D-05; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-05 - Multi-Agent and Subagent Orchestration.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-06 - AGI Memory Technology and Context Compaction: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-06 develops transferable capability in AGI Memory Technology and Context Compaction. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-06 - AGI Memory Technology and Context Compaction: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-06 - AGI Memory Technology and Context Compaction: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-06 - AGI Memory Technology and Context Compaction: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-06 - AGI Memory Technology and Context Compaction: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-06 - AGI Memory Technology and Context Compaction: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-06 - AGI Memory Technology and Context Compaction: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in AGI Memory Technology and Context Compaction is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-06 - AGI Memory Technology and Context Compaction: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-06

Trigger	Name the condition that activates D-06; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-06 - AGI Memory Technology and Context Compaction.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-06 - AGI Memory Technology and Context Compaction: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for AGI Memory Technology and Context Compaction. Hide some cases to reduce prompt overfitting. Measurement design for D-06 - AGI Memory Technology and Context Compaction: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-06 - AGI Memory Technology and Context Compaction: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-06 - AGI Memory Technology and Context Compaction: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in AGI Memory Technology and Context Compaction, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-06 - AGI Memory Technology and Context Compaction: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-06 - AGI Memory Technology and Context Compaction: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-06 - AGI Memory Technology and Context Compaction: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-06 - AGI Memory Technology and Context Compaction: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-06

Trigger	Name the condition that activates D-06; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-06 - AGI Memory Technology and Context Compaction.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-07 develops transferable capability in HyperKnowledge, KB-as-OS, and GraphRAG. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in HyperKnowledge, KB-as-OS, and GraphRAG is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-07

Trigger	Name the condition that activates D-07; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-07 - HyperKnowledge, KB-as-OS, and GraphRAG.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for HyperKnowledge, KB-as-OS, and GraphRAG. Hide some cases to reduce prompt overfitting. Target design for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in HyperKnowledge, KB-as-OS, and GraphRAG, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-07 - HyperKnowledge, KB-as-OS, and GraphRAG: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-07

Trigger	Name the condition that activates D-07; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-07 - HyperKnowledge, KB-as-OS, and GraphRAG.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-08 - Tool Use, Internet Research, Browser, and APIs: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-08 develops transferable capability in Tool Use, Internet Research, Browser, and APIs. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Governance analysis for D-08 - Tool Use, Internet Research, Browser, and APIs: connect domain objective to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Scenario design for D-08 - Tool Use, Internet Research, Browser, and APIs: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for capability contract. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Measurement design for D-08 - Tool Use, Internet Research, Browser, and APIs: assess failure taxonomy through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Target design for D-08 - Tool Use, Internet Research, Browser, and APIs: human role is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Architecture analysis for D-08 - Tool Use, Internet Research, Browser, and APIs: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Scenario design for D-08 - Tool Use, Internet Research, Browser, and APIs: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for knowledge and tools. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

TRANSFER REQUIREMENT

A capability in Tool Use, Internet Research, Browser, and APIs is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Scenario design for D-08 - Tool Use, Internet Research, Browser, and APIs: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for transfer requirement. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-08

Trigger	Name the condition that activates D-08; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-08 - Tool Use, Internet Research, Browser, and APIs.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-08 - Tool Use, Internet Research, Browser, and APIs: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Tool Use, Internet Research, Browser, and APIs. Hide some cases to reduce prompt overfitting. Governance analysis for D-08 - Tool Use, Internet Research, Browser, and APIs: connect benchmark pack to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Scenario design for D-08 - Tool Use, Internet Research, Browser, and APIs: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for metrics. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Target design for D-08 - Tool Use, Internet Research, Browser, and APIs: adversarial coverage is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Tool Use, Internet Research, Browser, and APIs, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Scenario design for D-08 - Tool Use, Internet Research, Browser, and APIs: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for open research. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Target design for D-08 - Tool Use, Internet Research, Browser, and APIs: baselines is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Scenario design for D-08 - Tool Use, Internet Research, Browser, and APIs: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for ablations. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Transfer analysis for D-08 - Tool Use, Internet Research, Browser, and APIs: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-08

Trigger	Name the condition that activates D-08; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-08 - Tool Use, Internet Research, Browser, and APIs.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-09 - Verification, Evidence, and Falsification Courts: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-09 develops transferable capability in Verification, Evidence, and Falsification Courts. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Transfer analysis for D-09 - Verification, Evidence, and Falsification Courts: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Target design for D-09 - Verification, Evidence, and Falsification Courts: capability contract is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Scenario design for D-09 - Verification, Evidence, and Falsification Courts: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for failure taxonomy. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Architecture analysis for D-09 - Verification, Evidence, and Falsification Courts: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Governance analysis for D-09 - Verification, Evidence, and Falsification Courts: connect twenty target pattern to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Target design for D-09 - Verification, Evidence, and Falsification Courts: knowledge and tools is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

TRANSFER REQUIREMENT

A capability in Verification, Evidence, and Falsification Courts is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Target design for D-09 - Verification, Evidence, and Falsification Courts: transfer requirement is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-09

Trigger	Name the condition that activates D-09; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-09 - Verification, Evidence, and Falsification Courts.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-09 - Verification, Evidence, and Falsification Courts: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Verification, Evidence, and Falsification Courts. Hide some cases to reduce prompt overfitting. Transfer analysis for D-09 - Verification, Evidence, and Falsification Courts: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Target design for D-09 - Verification, Evidence, and Falsification Courts: metrics is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Architecture analysis for D-09 - Verification, Evidence, and Falsification Courts: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Verification, Evidence, and Falsification Courts, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Target design for D-09 - Verification, Evidence, and Falsification Courts: open research is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Architecture analysis for D-09 - Verification, Evidence, and Falsification Courts: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Target design for D-09 - Verification, Evidence, and Falsification Courts: ablations is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Measurement design for D-09 - Verification, Evidence, and Falsification Courts: assess promotion decision through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-09

Trigger	Name the condition that activates D-09; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-09 - Verification, Evidence, and Falsification Courts.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-10 - Safety, Governance, Alignment, and Human Authority: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-10 develops transferable capability in Safety, Governance, Alignment, and Human Authority. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-10 - Safety, Governance, Alignment, and Human Authority: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-10 - Safety, Governance, Alignment, and Human Authority: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-10 - Safety, Governance, Alignment, and Human Authority: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-10 - Safety, Governance, Alignment, and Human Authority: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-10 - Safety, Governance, Alignment, and Human Authority: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-10 - Safety, Governance, Alignment, and Human Authority: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in Safety, Governance, Alignment, and Human Authority is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-10 - Safety, Governance, Alignment, and Human Authority: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-10

Trigger	Name the condition that activates D-10; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-10 - Safety, Governance, Alignment, and Human Authority.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-10 - Safety, Governance, Alignment, and Human Authority: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Safety, Governance, Alignment, and Human Authority. Hide some cases to reduce prompt overfitting. Target design for D-10 - Safety, Governance, Alignment, and Human Authority: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-10 - Safety, Governance, Alignment, and Human Authority: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-10 - Safety, Governance, Alignment, and Human Authority: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Safety, Governance, Alignment, and Human Authority, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-10 - Safety, Governance, Alignment, and Human Authority: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-10 - Safety, Governance, Alignment, and Human Authority: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-10 - Safety, Governance, Alignment, and Human Authority: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-10 - Safety, Governance, Alignment, and Human Authority: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-10

Trigger	Name the condition that activates D-10; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-10 - Safety, Governance, Alignment, and Human Authority.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-11 - Token, RPM, Latency, and Compute Economy: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-11 develops transferable capability in Token, RPM, Latency, and Compute Economy. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Architecture analysis for D-11 - Token, RPM, Latency, and Compute Economy: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Measurement design for D-11 - Token, RPM, Latency, and Compute Economy: assess capability contract through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Transfer analysis for D-11 - Token, RPM, Latency, and Compute Economy: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Scenario design for D-11 - Token, RPM, Latency, and Compute Economy: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for human role. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Target design for D-11 - Token, RPM, Latency, and Compute Economy: twenty target pattern is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Measurement design for D-11 - Token, RPM, Latency, and Compute Economy: assess knowledge and tools through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

TRANSFER REQUIREMENT

A capability in Token, RPM, Latency, and Compute Economy is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Measurement design for D-11 - Token, RPM, Latency, and Compute Economy: assess transfer requirement through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-11

Trigger	Name the condition that activates D-11; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-11 - Token, RPM, Latency, and Compute Economy.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-11 - Token, RPM, Latency, and Compute Economy: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Token, RPM, Latency, and Compute Economy. Hide some cases to reduce prompt overfitting. Architecture analysis for D-11 - Token, RPM, Latency, and Compute Economy: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Measurement design for D-11 - Token, RPM, Latency, and Compute Economy: assess metrics through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Scenario design for D-11 - Token, RPM, Latency, and Compute Economy: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for adversarial coverage. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Token, RPM, Latency, and Compute Economy, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Measurement design for D-11 - Token, RPM, Latency, and Compute Economy: assess open research through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Scenario design for D-11 - Token, RPM, Latency, and Compute Economy: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for baselines. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Measurement design for D-11 - Token, RPM, Latency, and Compute Economy: assess ablations through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Governance analysis for D-11 - Token, RPM, Latency, and Compute Economy: connect promotion decision to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-11

Trigger	Name the condition that activates D-11; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-11 - Token, RPM, Latency, and Compute Economy.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-12 develops transferable capability in Learning, HyperSkill Synthesis, and Governed Self-Improvement. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Transfer analysis for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Target design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: capability contract is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Scenario design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for failure taxonomy. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Architecture analysis for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Governance analysis for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: connect twenty target pattern to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Target design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: knowledge and tools is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

TRANSFER REQUIREMENT

A capability in Learning, HyperSkill Synthesis, and Governed Self-Improvement is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Target design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: transfer requirement is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-12

Trigger	Name the condition that activates D-12; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Learning, HyperSkill Synthesis, and Governed Self-Improvement. Hide some cases to reduce prompt overfitting. Transfer analysis for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Target design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: metrics is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Architecture analysis for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Learning, HyperSkill Synthesis, and Governed Self-Improvement, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Target design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: open research is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Architecture analysis for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Target design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: ablations is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Measurement design for D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement: assess promotion decision through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-12

Trigger	Name the condition that activates D-12; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-12 - Learning, HyperSkill Synthesis, and Governed Self-Improvement.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-13 - Software Architecture and Engineering: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-13 develops transferable capability in Software Architecture and Engineering. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Architecture analysis for D-13 - Software Architecture and Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Measurement design for D-13 - Software Architecture and Engineering: assess capability contract through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Transfer analysis for D-13 - Software Architecture and Engineering: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Scenario design for D-13 - Software Architecture and Engineering: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for human role. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Target design for D-13 - Software Architecture and Engineering: twenty target pattern is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Measurement design for D-13 - Software Architecture and Engineering: assess knowledge and tools through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

TRANSFER REQUIREMENT

A capability in Software Architecture and Engineering is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Measurement design for D-13 - Software Architecture and Engineering: assess transfer requirement through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX D-13

Trigger	Name the condition that activates D-13; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-13 - Software Architecture and Engineering.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-13 - Software Architecture and Engineering: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Software Architecture and Engineering. Hide some cases to reduce prompt overfitting. Architecture analysis for D-13 - Software Architecture and Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Measurement design for D-13 - Software Architecture and Engineering: assess metrics through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Scenario design for D-13 - Software Architecture and Engineering: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for adversarial coverage. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Software Architecture and Engineering, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Measurement design for D-13 - Software Architecture and Engineering: assess open research through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Scenario design for D-13 - Software Architecture and Engineering: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for baselines. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Measurement design for D-13 - Software Architecture and Engineering: assess ablations through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Governance analysis for D-13 - Software Architecture and Engineering: connect promotion decision to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-13

Trigger	Name the condition that activates D-13; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-13 - Software Architecture and Engineering.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-14 - Data Engineering, Analytics, and Machine Learning: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-14 develops transferable capability in Data Engineering, Analytics, and Machine Learning. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Architecture analysis for D-14 - Data Engineering, Analytics, and Machine Learning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Measurement design for D-14 - Data Engineering, Analytics, and Machine Learning: assess capability contract through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Transfer analysis for D-14 - Data Engineering, Analytics, and Machine Learning: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Scenario design for D-14 - Data Engineering, Analytics, and Machine Learning: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for human role. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Target design for D-14 - Data Engineering, Analytics, and Machine Learning: twenty target pattern is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Measurement design for D-14 - Data Engineering, Analytics, and Machine Learning: assess knowledge and tools through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

TRANSFER REQUIREMENT

A capability in Data Engineering, Analytics, and Machine Learning is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Measurement design for D-14 - Data Engineering, Analytics, and Machine Learning: assess transfer requirement through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-14

Trigger	Name the condition that activates D-14; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-14 - Data Engineering, Analytics, and Machine Learning.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-14 - Data Engineering, Analytics, and Machine Learning: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Data Engineering, Analytics, and Machine Learning. Hide some cases to reduce prompt overfitting. Architecture analysis for D-14 - Data Engineering, Analytics, and Machine Learning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Measurement design for D-14 - Data Engineering, Analytics, and Machine Learning: assess metrics through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Scenario design for D-14 - Data Engineering, Analytics, and Machine Learning: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for adversarial coverage. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Data Engineering, Analytics, and Machine Learning, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Measurement design for D-14 - Data Engineering, Analytics, and Machine Learning: assess open research through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Scenario design for D-14 - Data Engineering, Analytics, and Machine Learning: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for baselines. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Measurement design for D-14 - Data Engineering, Analytics, and Machine Learning: assess ablations through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Governance analysis for D-14 - Data Engineering, Analytics, and Machine Learning: connect promotion decision to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-14

Trigger	Name the condition that activates D-14; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-14 - Data Engineering, Analytics, and Machine Learning.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-15 - Cloud, DevOps, SRE, and Production Reliability: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-15 develops transferable capability in Cloud, DevOps, SRE, and Production Reliability. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Scenario design for D-15 - Cloud, DevOps, SRE, and Production Reliability: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for domain objective. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Governance analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: connect capability contract to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Architecture analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Transfer analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Measurement design for D-15 - Cloud, DevOps, SRE, and Production Reliability: assess twenty target pattern through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Governance analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: connect knowledge and tools to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

TRANSFER REQUIREMENT

A capability in Cloud, DevOps, SRE, and Production Reliability is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Governance analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: connect transfer requirement to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-15

Trigger	Name the condition that activates D-15; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-15 - Cloud, DevOps, SRE, and Production Reliability.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-15 - Cloud, DevOps, SRE, and Production Reliability: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Cloud, DevOps, SRE, and Production Reliability. Hide some cases to reduce prompt overfitting. Scenario design for D-15 - Cloud, DevOps, SRE, and Production Reliability: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for benchmark pack. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Governance analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: connect metrics to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Transfer analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Cloud, DevOps, SRE, and Production Reliability, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Governance analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: connect open research to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Transfer analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Governance analysis for D-15 - Cloud, DevOps, SRE, and Production Reliability: connect ablations to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Target design for D-15 - Cloud, DevOps, SRE, and Production Reliability: promotion decision is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-15

Trigger	Name the condition that activates D-15; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-15 - Cloud, DevOps, SRE, and Production Reliability.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-16 - Cybersecurity, Privacy, and Adversarial Resilience: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-16 develops transferable capability in Cybersecurity, Privacy, and Adversarial Resilience. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Architecture analysis for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Measurement design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: assess capability contract through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Transfer analysis for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Scenario design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for human role. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Target design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: twenty target pattern is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Measurement design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: assess knowledge and tools through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

TRANSFER REQUIREMENT

A capability in Cybersecurity, Privacy, and Adversarial Resilience is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Measurement design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: assess transfer requirement through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-16

Trigger	Name the condition that activates D-16; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-16 - Cybersecurity, Privacy, and Adversarial Resilience.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-16 - Cybersecurity, Privacy, and Adversarial Resilience: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Cybersecurity, Privacy, and Adversarial Resilience. Hide some cases to reduce prompt overfitting. Architecture analysis for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Measurement design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: assess metrics through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Scenario design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for adversarial coverage. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Cybersecurity, Privacy, and Adversarial Resilience, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Measurement design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: assess open research through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Scenario design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for baselines. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Measurement design for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: assess ablations through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Governance analysis for D-16 - Cybersecurity, Privacy, and Adversarial Resilience: connect promotion decision to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-16

Trigger	Name the condition that activates D-16; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-16 - Cybersecurity, Privacy, and Adversarial Resilience.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-17 - Research, Science, Mathematics, and Theory Building: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-17 develops transferable capability in Research, Science, Mathematics, and Theory Building. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-17 - Research, Science, Mathematics, and Theory Building: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-17 - Research, Science, Mathematics, and Theory Building: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-17 - Research, Science, Mathematics, and Theory Building: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-17 - Research, Science, Mathematics, and Theory Building: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-17 - Research, Science, Mathematics, and Theory Building: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-17 - Research, Science, Mathematics, and Theory Building: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in Research, Science, Mathematics, and Theory Building is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-17 - Research, Science, Mathematics, and Theory Building: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-17

Trigger	Name the condition that activates D-17; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-17 - Research, Science, Mathematics, and Theory Building.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-17 - Research, Science, Mathematics, and Theory Building: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Research, Science, Mathematics, and Theory Building. Hide some cases to reduce prompt overfitting. Measurement design for D-17 - Research, Science, Mathematics, and Theory Building: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-17 - Research, Science, Mathematics, and Theory Building: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-17 - Research, Science, Mathematics, and Theory Building: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Research, Science, Mathematics, and Theory Building, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-17 - Research, Science, Mathematics, and Theory Building: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-17 - Research, Science, Mathematics, and Theory Building: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-17 - Research, Science, Mathematics, and Theory Building: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-17 - Research, Science, Mathematics, and Theory Building: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-17

Trigger	Name the condition that activates D-17; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-17 - Research, Science, Mathematics, and Theory Building.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-18 - Documents, PDFs, Presentations, and Publishing: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-18 develops transferable capability in Documents, PDFs, Presentations, and Publishing. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Governance analysis for D-18 - Documents, PDFs, Presentations, and Publishing: connect domain objective to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Scenario design for D-18 - Documents, PDFs, Presentations, and Publishing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for capability contract. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Measurement design for D-18 - Documents, PDFs, Presentations, and Publishing: assess failure taxonomy through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Target design for D-18 - Documents, PDFs, Presentations, and Publishing: human role is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Architecture analysis for D-18 - Documents, PDFs, Presentations, and Publishing: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Scenario design for D-18 - Documents, PDFs, Presentations, and Publishing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for knowledge and tools. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

TRANSFER REQUIREMENT

A capability in Documents, PDFs, Presentations, and Publishing is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Scenario design for D-18 - Documents, PDFs, Presentations, and Publishing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for transfer requirement. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-18

Trigger	Name the condition that activates D-18; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-18 - Documents, PDFs, Presentations, and Publishing.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-18 - Documents, PDFs, Presentations, and Publishing: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Documents, PDFs, Presentations, and Publishing. Hide some cases to reduce prompt overfitting. Governance analysis for D-18 - Documents, PDFs, Presentations, and Publishing: connect benchmark pack to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Scenario design for D-18 - Documents, PDFs, Presentations, and Publishing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for metrics. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Target design for D-18 - Documents, PDFs, Presentations, and Publishing: adversarial coverage is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Documents, PDFs, Presentations, and Publishing, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Scenario design for D-18 - Documents, PDFs, Presentations, and Publishing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for open research. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Target design for D-18 - Documents, PDFs, Presentations, and Publishing: baselines is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Scenario design for D-18 - Documents, PDFs, Presentations, and Publishing: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for ablations. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Transfer analysis for D-18 - Documents, PDFs, Presentations, and Publishing: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-18

Trigger	Name the condition that activates D-18; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-18 - Documents, PDFs, Presentations, and Publishing.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-19 - Language, Translation, Localization, and Cultural Reasoning: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-19 develops transferable capability in Language, Translation, Localization, and Cultural Reasoning. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-19 - Language, Translation, Localization, and Cultural Reasoning: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-19 - Language, Translation, Localization, and Cultural Reasoning: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in Language, Translation, Localization, and Cultural Reasoning is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-19

Trigger	Name the condition that activates D-19; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-19 - Language, Translation, Localization, and Cultural Reasoning.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-19 - Language, Translation, Localization, and Cultural Reasoning: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Language, Translation, Localization, and Cultural Reasoning. Hide some cases to reduce prompt overfitting. Measurement design for D-19 - Language, Translation, Localization, and Cultural Reasoning: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Language, Translation, Localization, and Cultural Reasoning, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-19 - Language, Translation, Localization, and Cultural Reasoning: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-19 - Language, Translation, Localization, and Cultural Reasoning: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-19

Trigger	Name the condition that activates D-19; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-19 - Language, Translation, Localization, and Cultural Reasoning.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-20 - Health, Medicine, Psychology, and Biomedical Information: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-20 develops transferable capability in Health, Medicine, Psychology, and Biomedical Information. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-20 - Health, Medicine, Psychology, and Biomedical Information: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-20 - Health, Medicine, Psychology, and Biomedical Information: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-20 - Health, Medicine, Psychology, and Biomedical Information: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in Health, Medicine, Psychology, and Biomedical Information is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-20

Trigger	Name the condition that activates D-20; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-20 - Health, Medicine, Psychology, and Biomedical Information.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-20 - Health, Medicine, Psychology, and Biomedical Information: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Health, Medicine, Psychology, and Biomedical Information. Hide some cases to reduce prompt overfitting. Target design for D-20 - Health, Medicine, Psychology, and Biomedical Information: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-20 - Health, Medicine, Psychology, and Biomedical Information: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Health, Medicine, Psychology, and Biomedical Information, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-20 - Health, Medicine, Psychology, and Biomedical Information: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-20 - Health, Medicine, Psychology, and Biomedical Information: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-20

Trigger	Name the condition that activates D-20; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-20 - Health, Medicine, Psychology, and Biomedical Information.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-21 - Multimodal Media, Visual Design, Audio, and Video: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-21 develops transferable capability in Multimodal Media, Visual Design, Audio, and Video. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Governance analysis for D-21 - Multimodal Media, Visual Design, Audio, and Video: connect domain objective to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Scenario design for D-21 - Multimodal Media, Visual Design, Audio, and Video: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for capability contract. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Measurement design for D-21 - Multimodal Media, Visual Design, Audio, and Video: assess failure taxonomy through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Target design for D-21 - Multimodal Media, Visual Design, Audio, and Video: human role is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Architecture analysis for D-21 - Multimodal Media, Visual Design, Audio, and Video: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Scenario design for D-21 - Multimodal Media, Visual Design, Audio, and Video: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for knowledge and tools. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

TRANSFER REQUIREMENT

A capability in Multimodal Media, Visual Design, Audio, and Video is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Scenario design for D-21 - Multimodal Media, Visual Design, Audio, and Video: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for transfer requirement. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-21

Trigger	Name the condition that activates D-21; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-21 - Multimodal Media, Visual Design, Audio, and Video.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-21 - Multimodal Media, Visual Design, Audio, and Video: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Multimodal Media, Visual Design, Audio, and Video. Hide some cases to reduce prompt overfitting. Governance analysis for D-21 - Multimodal Media, Visual Design, Audio, and Video: connect benchmark pack to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Scenario design for D-21 - Multimodal Media, Visual Design, Audio, and Video: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for metrics. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Target design for D-21 - Multimodal Media, Visual Design, Audio, and Video: adversarial coverage is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Multimodal Media, Visual Design, Audio, and Video, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Scenario design for D-21 - Multimodal Media, Visual Design, Audio, and Video: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for open research. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Target design for D-21 - Multimodal Media, Visual Design, Audio, and Video: baselines is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Scenario design for D-21 - Multimodal Media, Visual Design, Audio, and Video: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for ablations. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Transfer analysis for D-21 - Multimodal Media, Visual Design, Audio, and Video: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-21

Trigger	Name the condition that activates D-21; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-21 - Multimodal Media, Visual Design, Audio, and Video.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-22 - Business, Product, Strategy, and Value Creation: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-22 develops transferable capability in Business, Product, Strategy, and Value Creation. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-22 - Business, Product, Strategy, and Value Creation: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-22 - Business, Product, Strategy, and Value Creation: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-22 - Business, Product, Strategy, and Value Creation: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-22 - Business, Product, Strategy, and Value Creation: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-22 - Business, Product, Strategy, and Value Creation: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-22 - Business, Product, Strategy, and Value Creation: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in Business, Product, Strategy, and Value Creation is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-22 - Business, Product, Strategy, and Value Creation: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-22

Trigger	Name the condition that activates D-22; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-22 - Business, Product, Strategy, and Value Creation.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-22 - Business, Product, Strategy, and Value Creation: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Business, Product, Strategy, and Value Creation. Hide some cases to reduce prompt overfitting. Target design for D-22 - Business, Product, Strategy, and Value Creation: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-22 - Business, Product, Strategy, and Value Creation: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-22 - Business, Product, Strategy, and Value Creation: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Business, Product, Strategy, and Value Creation, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-22 - Business, Product, Strategy, and Value Creation: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-22 - Business, Product, Strategy, and Value Creation: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-22 - Business, Product, Strategy, and Value Creation: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-22 - Business, Product, Strategy, and Value Creation: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-22

Trigger	Name the condition that activates D-22; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-22 - Business, Product, Strategy, and Value Creation.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-23 - Ethical Marketing, Sales, Outreach, and Growth: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-23 develops transferable capability in Ethical Marketing, Sales, Outreach, and Growth. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Scenario design for D-23 - Ethical Marketing, Sales, Outreach, and Growth: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for domain objective. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Governance analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: connect capability contract to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Architecture analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Transfer analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Measurement design for D-23 - Ethical Marketing, Sales, Outreach, and Growth: assess twenty target pattern through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Governance analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: connect knowledge and tools to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

TRANSFER REQUIREMENT

A capability in Ethical Marketing, Sales, Outreach, and Growth is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Governance analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: connect transfer requirement to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-23

Trigger	Name the condition that activates D-23; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-23 - Ethical Marketing, Sales, Outreach, and Growth.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-23 - Ethical Marketing, Sales, Outreach, and Growth: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Ethical Marketing, Sales, Outreach, and Growth. Hide some cases to reduce prompt overfitting. Scenario design for D-23 - Ethical Marketing, Sales, Outreach, and Growth: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for benchmark pack. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Governance analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: connect metrics to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Transfer analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Ethical Marketing, Sales, Outreach, and Growth, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Governance analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: connect open research to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Transfer analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Governance analysis for D-23 - Ethical Marketing, Sales, Outreach, and Growth: connect ablations to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Target design for D-23 - Ethical Marketing, Sales, Outreach, and Growth: promotion decision is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-23

Trigger	Name the condition that activates D-23; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-23 - Ethical Marketing, Sales, Outreach, and Growth.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-24 - Finance, Economics, Risk, and Decision Analysis: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-24 develops transferable capability in Finance, Economics, Risk, and Decision Analysis. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Transfer analysis for D-24 - Finance, Economics, Risk, and Decision Analysis: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Target design for D-24 - Finance, Economics, Risk, and Decision Analysis: capability contract is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Scenario design for D-24 - Finance, Economics, Risk, and Decision Analysis: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for failure taxonomy. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Architecture analysis for D-24 - Finance, Economics, Risk, and Decision Analysis: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Governance analysis for D-24 - Finance, Economics, Risk, and Decision Analysis: connect twenty target pattern to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Target design for D-24 - Finance, Economics, Risk, and Decision Analysis: knowledge and tools is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

TRANSFER REQUIREMENT

A capability in Finance, Economics, Risk, and Decision Analysis is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Target design for D-24 - Finance, Economics, Risk, and Decision Analysis: transfer requirement is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-24

Trigger	Name the condition that activates D-24; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-24 - Finance, Economics, Risk, and Decision Analysis.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-24 - Finance, Economics, Risk, and Decision Analysis: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Finance, Economics, Risk, and Decision Analysis. Hide some cases to reduce prompt overfitting. Transfer analysis for D-24 - Finance, Economics, Risk, and Decision Analysis: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Target design for D-24 - Finance, Economics, Risk, and Decision Analysis: metrics is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Architecture analysis for D-24 - Finance, Economics, Risk, and Decision Analysis: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Finance, Economics, Risk, and Decision Analysis, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Target design for D-24 - Finance, Economics, Risk, and Decision Analysis: open research is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Architecture analysis for D-24 - Finance, Economics, Risk, and Decision Analysis: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Target design for D-24 - Finance, Economics, Risk, and Decision Analysis: ablations is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Measurement design for D-24 - Finance, Economics, Risk, and Decision Analysis: assess promotion decision through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-24

Trigger	Name the condition that activates D-24; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-24 - Finance, Economics, Risk, and Decision Analysis.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-25 - Legal Information, Policy, and Compliance Engineering: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-25 develops transferable capability in Legal Information, Policy, and Compliance Engineering. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-25 - Legal Information, Policy, and Compliance Engineering: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-25 - Legal Information, Policy, and Compliance Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-25 - Legal Information, Policy, and Compliance Engineering: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-25 - Legal Information, Policy, and Compliance Engineering: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-25 - Legal Information, Policy, and Compliance Engineering: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-25 - Legal Information, Policy, and Compliance Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in Legal Information, Policy, and Compliance Engineering is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-25 - Legal Information, Policy, and Compliance Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-25

Trigger	Name the condition that activates D-25; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-25 - Legal Information, Policy, and Compliance Engineering.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-25 - Legal Information, Policy, and Compliance Engineering: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Legal Information, Policy, and Compliance Engineering. Hide some cases to reduce prompt overfitting. Measurement design for D-25 - Legal Information, Policy, and Compliance Engineering: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-25 - Legal Information, Policy, and Compliance Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-25 - Legal Information, Policy, and Compliance Engineering: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Legal Information, Policy, and Compliance Engineering, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-25 - Legal Information, Policy, and Compliance Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-25 - Legal Information, Policy, and Compliance Engineering: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-25 - Legal Information, Policy, and Compliance Engineering: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-25 - Legal Information, Policy, and Compliance Engineering: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-25

Trigger	Name the condition that activates D-25; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-25 - Legal Information, Policy, and Compliance Engineering.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-26 - Human Collaboration, Communication, and Teaching: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-26 develops transferable capability in Human Collaboration, Communication, and Teaching. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Scenario design for D-26 - Human Collaboration, Communication, and Teaching: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for domain objective. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Governance analysis for D-26 - Human Collaboration, Communication, and Teaching: connect capability contract to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Architecture analysis for D-26 - Human Collaboration, Communication, and Teaching: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Transfer analysis for D-26 - Human Collaboration, Communication, and Teaching: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Measurement design for D-26 - Human Collaboration, Communication, and Teaching: assess twenty target pattern through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Governance analysis for D-26 - Human Collaboration, Communication, and Teaching: connect knowledge and tools to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

TRANSFER REQUIREMENT

A capability in Human Collaboration, Communication, and Teaching is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Governance analysis for D-26 - Human Collaboration, Communication, and Teaching: connect transfer requirement to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPERATIONAL EVIDENCE MATRIX

D-26

Trigger	Name the condition that activates D-26; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-26 - Human Collaboration, Communication, and Teaching.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-26 - Human Collaboration, Communication, and Teaching: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Human Collaboration, Communication, and Teaching. Hide some cases to reduce prompt overfitting. Scenario design for D-26 - Human Collaboration, Communication, and Teaching: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for benchmark pack. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Governance analysis for D-26 - Human Collaboration, Communication, and Teaching: connect metrics to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Transfer analysis for D-26 - Human Collaboration, Communication, and Teaching: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Human Collaboration, Communication, and Teaching, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Governance analysis for D-26 - Human Collaboration, Communication, and Teaching: connect open research to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Transfer analysis for D-26 - Human Collaboration, Communication, and Teaching: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Governance analysis for D-26 - Human Collaboration, Communication, and Teaching: connect ablations to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Target design for D-26 - Human Collaboration, Communication, and Teaching: promotion decision is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-26

Trigger	Name the condition that activates D-26; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-26 - Human Collaboration, Communication, and Teaching.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-27 - Robotics, Embodied Agency, and World Models: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-27 develops transferable capability in Robotics, Embodied Agency, and World Models. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Governance analysis for D-27 - Robotics, Embodied Agency, and World Models: connect domain objective to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Scenario design for D-27 - Robotics, Embodied Agency, and World Models: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for capability contract. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Measurement design for D-27 - Robotics, Embodied Agency, and World Models: assess failure taxonomy through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Target design for D-27 - Robotics, Embodied Agency, and World Models: human role is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Architecture analysis for D-27 - Robotics, Embodied Agency, and World Models: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Scenario design for D-27 - Robotics, Embodied Agency, and World Models: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for knowledge and tools. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

TRANSFER REQUIREMENT

A capability in Robotics, Embodied Agency, and World Models is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Scenario design for D-27 - Robotics, Embodied Agency, and World Models: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for transfer requirement. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-27

Trigger	Name the condition that activates D-27; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-27 - Robotics, Embodied Agency, and World Models.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-27 - Robotics, Embodied Agency, and World Models: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Robotics, Embodied Agency, and World Models. Hide some cases to reduce prompt overfitting. Governance analysis for D-27 - Robotics, Embodied Agency, and World Models: connect benchmark pack to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Scenario design for D-27 - Robotics, Embodied Agency, and World Models: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for metrics. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Target design for D-27 - Robotics, Embodied Agency, and World Models: adversarial coverage is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Robotics, Embodied Agency, and World Models, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Scenario design for D-27 - Robotics, Embodied Agency, and World Models: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for open research. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Target design for D-27 - Robotics, Embodied Agency, and World Models: baselines is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Scenario design for D-27 - Robotics, Embodied Agency, and World Models: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for ablations. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Transfer analysis for D-27 - Robotics, Embodied Agency, and World Models: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-27

Trigger	Name the condition that activates D-27; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-27 - Robotics, Embodied Agency, and World Models.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-28 - Causal Inference, Simulation, Forecasting, and Optimization: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-28 develops transferable capability in Causal Inference, Simulation, Forecasting, and Optimization. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Transfer analysis for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Target design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: capability contract is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Scenario design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for failure taxonomy. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Architecture analysis for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Governance analysis for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: connect twenty target pattern to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Target design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: knowledge and tools is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

TRANSFER REQUIREMENT

A capability in Causal Inference, Simulation, Forecasting, and Optimization is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Target design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: transfer requirement is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

OPERATIONAL EVIDENCE MATRIX

D-28

Trigger	Name the condition that activates D-28; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-28 - Causal Inference, Simulation, Forecasting, and Optimization.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-28 - Causal Inference, Simulation, Forecasting, and Optimization: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Causal Inference, Simulation, Forecasting, and Optimization. Hide some cases to reduce prompt overfitting. Transfer analysis for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Target design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: metrics is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Architecture analysis for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Causal Inference, Simulation, Forecasting, and Optimization, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Target design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: open research is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Architecture analysis for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Target design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: ablations is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Measurement design for D-28 - Causal Inference, Simulation, Forecasting, and Optimization: assess promotion decision through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPERATIONAL EVIDENCE MATRIX

D-28

Trigger	Name the condition that activates D-28; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-28 - Causal Inference, Simulation, Forecasting, and Optimization.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-29 - Agentic Deployment, Integration, and Observability: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-29 develops transferable capability in Agentic Deployment, Integration, and Observability. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Target design for D-29 - Agentic Deployment, Integration, and Observability: domain objective is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Transfer analysis for D-29 - Agentic Deployment, Integration, and Observability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Governance analysis for D-29 - Agentic Deployment, Integration, and Observability: connect failure taxonomy to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Measurement design for D-29 - Agentic Deployment, Integration, and Observability: assess human role through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Scenario design for D-29 - Agentic Deployment, Integration, and Observability: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for twenty target pattern. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Transfer analysis for D-29 - Agentic Deployment, Integration, and Observability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

TRANSFER REQUIREMENT

A capability in Agentic Deployment, Integration, and Observability is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Transfer analysis for D-29 - Agentic Deployment, Integration, and Observability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

OPERATIONAL EVIDENCE MATRIX

D-29

Trigger	Name the condition that activates D-29; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-29 - Agentic Deployment, Integration, and Observability.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-29 - Agentic Deployment, Integration, and Observability: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for Agentic Deployment, Integration, and Observability. Hide some cases to reduce prompt overfitting. Target design for D-29 - Agentic Deployment, Integration, and Observability: benchmark pack is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Transfer analysis for D-29 - Agentic Deployment, Integration, and Observability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Measurement design for D-29 - Agentic Deployment, Integration, and Observability: assess adversarial coverage through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in Agentic Deployment, Integration, and Observability, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Transfer analysis for D-29 - Agentic Deployment, Integration, and Observability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Measurement design for D-29 - Agentic Deployment, Integration, and Observability: assess baselines through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Transfer analysis for D-29 - Agentic Deployment, Integration, and Observability: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Architecture analysis for D-29 - Agentic Deployment, Integration, and Observability: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-29

Trigger	Name the condition that activates D-29; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-29 - Agentic Deployment, Integration, and Observability.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: Capability Scope and Architecture

DOMAIN OBJECTIVE

D-30 develops transferable capability in MICAGFI Neo-AGI Meta-Architecture and Benchmarks. The domain is a test bed for the kernel, not a claim that one architecture already solves the entire field. Measurement design for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: assess domain objective through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

CAPABILITY CONTRACT

Define task, expected artifact, evidence threshold, authority, budget, risk, stop rules, world-state delta and residual uncertainty before execution. Architecture analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

FAILURE TAXONOMY

Track misunderstanding, unsupported claim, wrong route, stale data, tool error, authority violation, unreconciled state, hidden cost, weak calibration and failed recovery. Target design for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: failure taxonomy is decomposed into observable tasks with declared inputs, output artifact, evidence threshold, permitted tools, time horizon and risk. Include easy, representative and bottleneck cases rather than only spectacular demonstrations. Each target has a negative activation example and a human-review boundary. Coverage is reported by capability and failure class so average performance cannot conceal an untested critical pathway.

HUMAN ROLE

Human authority sets goals, irreversible boundaries and acceptable risk. Evaluation includes comprehension, correction burden, option value and the ability to revoke control. Governance analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: connect human role to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

TWENTY TARGET PATTERN

Targets span framing, acquisition, transformation, generation, tool use, verification, governance, recovery and transfer. Each target has positive, negative, edge, adversarial and constrained cases. Transfer analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: change wording, data source, model provider, tool interface, user profile and operational constraint while preserving the underlying task. Distinguish adaptation from manual reconstruction. Measure retained capability, new failure signatures and additional supervision. A target that passes only on its construction template is not evidence of generalization. Negative transfer is catalogued and used to refine routing and boundary conditions.

KNOWLEDGE AND TOOLS

Route the smallest relevant HyperKnowledge set and one primary HyperSkill. Add tools only when they create observable value. Every external action has a typed contract and postcondition. Architecture analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

TRANSFER REQUIREMENT

A capability in MICAGFI Neo-AGI Meta-Architecture and Benchmarks is not general until it transfers across task wording, models, datasets, tools and environmental constraints without manual reconstruction. Architecture analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

OPERATIONAL EVIDENCE MATRIX

D-30

Trigger	Name the condition that activates D-30; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: Benchmark and Research Agenda

BENCHMARK PACK

Build clean, ambiguous, adversarial, long-horizon, constrained and transfer scenarios for MICAGFI Neo-AGI Meta-Architecture and Benchmarks. Hide some cases to reduce prompt overfitting. Measurement design for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: assess benchmark pack through artifact correctness, evidence support, calibration, cost, latency, route sparsity, tool reliability, authority adherence, external-effect accuracy, recovery and human correction burden. Report distributions, confidence intervals and worst-case results where possible. Multi-objective tradeoffs remain visible; a cheaper route that fails critical safety is not efficient, while an expensive mechanism with no marginal gain is not intelligent depth.

METRICS

Score artifact correctness, evidence quality, calibration, cost, latency, route sparsity, tool reliability, external-effect accuracy, recovery and human correction. Architecture analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

ADVERSARIAL COVERAGE

Inject conflicting instructions, stale evidence, misleading sources, broken APIs, context pressure, rate limits, permission ambiguity and partial external mutations. Governance analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: connect adversarial coverage to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

OPEN RESEARCH

Identify which mechanism creates the largest causal difference in MICAGFI Neo-AGI Meta-Architecture and Benchmarks, which failures remain domain-specific, and whether improvements transfer without increasing governance risk. Architecture analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

هذا domain هو test bed للـkernel، وليس ادعاء بأن المجال حل بالكامل.

BASELINES

Use ordinary prompting, a competent single-agent workflow, fixed model and tools, no-memory mode, no-court mode and an expert rule where appropriate. Governance analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: connect baselines to authority, privacy, domain regulation, evidence obligations and human escalation. Consequential actions declare reversibility and approval requirements. Sensitive records use minimization, purpose limitation and retention rules. Evaluation includes false confidence and inappropriate action, not only answer accuracy. The user must be able to inspect, correct, pause and revoke the route without losing the evidence needed for recovery.

ABLATIONS

Remove question governance, sparse routing, memory, world model, courts, subagents and rollback separately. Record when simplification improves results. Architecture analysis for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: identify which work belongs to the model, retrieval, deterministic code, specialist agent, external tool, court and human. Use the smallest sufficient combination and record why each component is present. Compare serial and parallel plans, measure coordination loss and test graceful degradation when a component is absent. Domain value comes from verified composition, not from routing every task to the largest model.

PROMOTION DECISION

ADOPT only when matched evidence supports advantage and critical gates pass. REVISE for mixed results; HOLD for missing evidence; ROLLBACK after regression or authority failure. Scenario design for D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks: generate clean, ambiguous, constrained, adversarial, long-horizon and transfer variants for promotion decision. Hidden variants reduce prompt memorization. Preserve realistic dependency failure, stale evidence, inconsistent instructions and incomplete external state. Scenario metadata records difficulty factors without leaking the answer to the agent. A domain specialist reviews whether the cases measure genuine capability rather than superficial terminology matching.

OPERATIONAL EVIDENCE MATRIX

D-30

Trigger	Name the condition that activates D-30; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by D-30 - MICAGFI Neo-AGI Meta-Architecture and Benchmarks.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Constitutional Identity - Skills 1 to 8

CATEGORY ROLE

Constitutional Identity contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-001 CONSTITUTIONAL-IDENTITY-IDENTITY-MANIFEST

MH-001 - constitutional-identity-identity-manifest. Execute identity manifest as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-003 CONSTITUTIONAL-IDENTITY-CONSTITUTIONAL-RESOLUTION

MH-003 - constitutional-identity-constitutional-resolution. Execute constitutional resolution as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-005 CONSTITUTIONAL-IDENTITY-AUTHORIZATION-BOUNDARIES

MH-005 - constitutional-identity-authorization-boundaries. Execute authorization boundaries as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-007 CONSTITUTIONAL-IDENTITY-CLAIM-CALIBRATION

MH-007 - constitutional-identity-claim-calibration. Execute claim calibration as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-002 CONSTITUTIONAL-IDENTITY-MISSION-ALIGNMENT

MH-002 - constitutional-identity-mission-alignment. Execute mission alignment as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-004 CONSTITUTIONAL-IDENTITY-VALUE-CONFLICT-ARBITRATION

MH-004 - constitutional-identity-value-conflict-arbitration. Execute value conflict arbitration as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-006 CONSTITUTIONAL-IDENTITY-EPISTEMIC-HUMILITY

MH-006 - constitutional-identity-epistemic-humility. Execute epistemic humility as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-008 CONSTITUTIONAL-IDENTITY-INTENT-PRESERVATION

MH-008 - constitutional-identity-intent-preservation. Execute intent preservation as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

CONSTITUTIONAL_IDENTITY

Trigger	Name the condition that activates CONSTITUTIONAL_IDENTITY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Constitutional Identity - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Constitutional Identity - Skills 9 to 15

CATEGORY ROLE

Constitutional Identity contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-009 CONSTITUTIONAL-IDENTITY-LONG-HORIZON-GOALS

MH-009 - constitutional-identity-long-horizon-goals. Execute long horizon goals as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-011 CONSTITUTIONAL-IDENTITY-SCOPE-DEFINITION

MH-011 - constitutional-identity-scope-definition. Execute scope definition as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-013 CONSTITUTIONAL-IDENTITY-IDENTITY-DRIFT-DETECTION

MH-013 - constitutional-identity-identity-drift-detection. Execute identity drift detection as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-015 CONSTITUTIONAL-IDENTITY-PURPOSE-TO-ACTION

MH-015 - constitutional-identity-purpose-to-action. Execute purpose to action as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-010 CONSTITUTIONAL-IDENTITY-NON-DECEPTION

MH-010 - constitutional-identity-non-deception. Execute non deception as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-012 CONSTITUTIONAL-IDENTITY-AGENT-ROLE-INSTANTIATION

MH-012 - constitutional-identity-agent-role-instantiation. Execute agent role instantiation as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-014 CONSTITUTIONAL-IDENTITY-CONSTITUTIONAL-EXCEPTIONS

MH-014 - constitutional-identity-constitutional-exceptions. Execute constitutional exceptions as an auditable capability within the Constitutional Identity and Agency layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

CONSTITUTIONAL_IDENTITY

Trigger	Name the condition that activates CONSTITUTIONAL_IDENTITY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Constitutional Identity - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Fattah Reasoning - Skills 1 to 8

CATEGORY ROLE

Fattah Reasoning contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-016 FATTAH-REASONING-PROBLEM-OPENING

MH-016 - fattah-reasoning-problem-opening. Execute problem opening as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-018 FATTAH-REASONING-SYNTHETIC-COMPRESSION

MH-018 - fattah-reasoning-synthetic-compression. Execute synthetic compression as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-020 FATTAH-REASONING-INNER-TIME-SCENARIOS

MH-020 - fattah-reasoning-inner-time-scenarios. Execute inner time scenarios as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-022 FATTAH-REASONING-CAUSAL-DECOMPOSITION

MH-022 - fattah-reasoning-causal-decomposition. Execute causal decomposition as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-017 FATTAH-REASONING-DIFFERENCE-EXTRACTION

MH-017 - fattah-reasoning-difference-extraction. Execute difference extraction as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-019 FATTAH-REASONING-META-COMPLEXITY-MAPPING

MH-019 - fattah-reasoning-meta-complexity-mapping. Execute meta complexity mapping as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-021 FATTAH-REASONING-ALTERNATIVE-SPACE-GENERATION

MH-021 - fattah-reasoning-alternative-space-generation. Execute alternative space generation as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-023 FATTAH-REASONING-ASSUMPTION-SURFACING

MH-023 - fattah-reasoning-assumption-surfacing. Execute assumption surfacing as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

FATTAH_REASONING

Trigger	Name the condition that activates FATTAH_REASONING; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Fattah Reasoning - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Fattah Reasoning - Skills 9 to 15

CATEGORY ROLE

Fattah Reasoning contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-024 FATTAH-REASONING-HIDDEN-VARIABLE-SEARCH

MH-024 - fattah-reasoning-hidden-variable-search. Execute hidden variable search as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-026 FATTAH-REASONING-THEORY-TO-MECHANISM

MH-026 - fattah-reasoning-theory-to-mechanism. Execute theory to mechanism as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-028 FATTAH-REASONING-MULTI-SCALE-ABSTRACTION

MH-028 - fattah-reasoning-multi-scale-abstraction. Execute multi scale abstraction as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-030 FATTAH-REASONING-REASONING-DEPTH

MH-030 - fattah-reasoning-reasoning-depth. Execute reasoning depth as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-025 FATTAH-REASONING-CROSS-DOMAIN-ANALOGY

MH-025 - fattah-reasoning-cross-domain-analogy. Execute cross domain analogy as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-027 FATTAH-REASONING-CONSTRAINT-REVERSAL

MH-027 - fattah-reasoning-constraint-reversal. Execute constraint reversal as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-029 FATTAH-REASONING-NOVEL-HYPOTHESIS

MH-029 - fattah-reasoning-novel-hypothesis. Execute novel hypothesis as an auditable capability within the AGI Fattah Reasoning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

FATTAH_REASONING

Trigger	Name the condition that activates FATTAH_REASONING; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Fattah Reasoning - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Gap Filling Evidence - Skills 1 to 8

CATEGORY ROLE

Gap Filling Evidence contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-031 GAP-FILLING-EVIDENCE-DEFINITIONS-GAP

MH-031 - gap-filling-evidence-definitions-gap. Execute definitions gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-033 GAP-FILLING-EVIDENCE-EVIDENCE-GAP

MH-033 - gap-filling-evidence-evidence-gap. Execute evidence gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-035 GAP-FILLING-EVIDENCE-MEASUREMENT-GAP

MH-035 - gap-filling-evidence-measurement-gap. Execute measurement gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-037 GAP-FILLING-EVIDENCE-VALIDATION-GAP

MH-037 - gap-filling-evidence-validation-gap. Execute validation gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-032 GAP-FILLING-EVIDENCE-BOUNDARY-GAP

MH-032 - gap-filling-evidence-boundary-gap. Execute boundary gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-034 GAP-FILLING-EVIDENCE-MECHANISM-GAP

MH-034 - gap-filling-evidence-mechanism-gap. Execute mechanism gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-036 GAP-FILLING-EVIDENCE-IMPLEMENTATION-GAP

MH-036 - gap-filling-evidence-implementation-gap. Execute implementation gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-038 GAP-FILLING-EVIDENCE-ETHICS-GAP

MH-038 - gap-filling-evidence-ethics-gap. Execute ethics gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

GAP_FILLING_EVIDENCE

Trigger	Name the condition that activates GAP_FILLING_EVIDENCE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Gap Filling Evidence - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Gap Filling Evidence - Skills 9 to 15

CATEGORY ROLE

Gap Filling Evidence contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-039 GAP-FILLING-EVIDENCE-BUSINESS-GAP

MH-039 - gap-filling-evidence-business-gap. Execute business gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-041 GAP-FILLING-EVIDENCE-EVIDENCE-GRADING

MH-041 - gap-filling-evidence-evidence-grading. Execute evidence grading as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-043 GAP-FILLING-EVIDENCE-CONTRADICTION-COURT

MH-043 - gap-filling-evidence-contradiction-court. Execute contradiction court as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-045 GAP-FILLING-EVIDENCE-COMPLETENESS-VERDICT

MH-045 - gap-filling-evidence-completeness-verdict. Execute completeness verdict as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-040 GAP-FILLING-EVIDENCE-COMMUNICATION-GAP

MH-040 - gap-filling-evidence-communication-gap. Execute communication gap as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-042 GAP-FILLING-EVIDENCE-CLAIM-SOURCE-LINKING

MH-042 - gap-filling-evidence-claim-source-linking. Execute claim source linking as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-044 GAP-FILLING-EVIDENCE-COUNTEREXAMPLE-COURT

MH-044 - gap-filling-evidence-counterexample-court. Execute counterexample court as an auditable capability within the Gap Filling and Evidence Court layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=low. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

GAP_FILLING_EVIDENCE

Trigger	Name the condition that activates GAP_FILLING_EVIDENCE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Gap Filling Evidence - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Project World Model - Skills 1 to 8

CATEGORY ROLE

Project World Model contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-046 PROJECT-WORLD-MODEL-PROJECT-DISCOVERY

MH-046 - project-world-model-project-discovery. Execute project discovery as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-048 PROJECT-WORLD-MODEL-STATE-SYNCHRONIZATION

MH-048 - project-world-model-state-synchronization. Execute state synchronization as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-050 PROJECT-WORLD-MODEL-CANONICAL-URL-REGISTRY

MH-050 - project-world-model-canonical-url-registry. Execute canonical url registry as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-052 PROJECT-WORLD-MODEL-DEPLOYMENT-INVENTORY

MH-052 - project-world-model-deployment-inventory. Execute deployment inventory as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-047 PROJECT-WORLD-MODEL-PROJECT-CANONICALIZATION

MH-047 - project-world-model-project-canonicalization. Execute project canonicalization as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-049 PROJECT-WORLD-MODEL-RELATIONSHIP-MAPPING

MH-049 - project-world-model-relationship-mapping. Execute relationship mapping as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-051 PROJECT-WORLD-MODEL-REPOSITORY-REGISTRY

MH-051 - project-world-model-repository-registry. Execute repository registry as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-053 PROJECT-WORLD-MODEL-ASSET-PROVENANCE

MH-053 - project-world-model-asset-provenance. Execute asset provenance as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

PROJECT_WORLD_MODEL

Trigger	Name the condition that activates PROJECT_WORLD_MODEL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Project World Model - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Project World Model - Skills 9 to 15

CATEGORY ROLE

Project World Model contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-054 PROJECT-WORLD-MODEL-OBJECTIVE-LINKING

MH-054 - project-world-model-objective-linking. Execute objective linking as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-056 PROJECT-WORLD-MODEL-OWNER-RESOLUTION

MH-056 - project-world-model-owner-resolution. Execute owner resolution as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-058 PROJECT-WORLD-MODEL-ARCHIVE-CLASSIFICATION

MH-058 - project-world-model-archive-classification. Execute archive classification as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-060 PROJECT-WORLD-MODEL-CONTEXT-PACKAGING

MH-060 - project-world-model-context-packaging. Execute context packaging as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-055 PROJECT-WORLD-MODEL-DEPENDENCY-GRAPH

MH-055 - project-world-model-dependency-graph. Execute dependency graph as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-057 PROJECT-WORLD-MODEL-PROJECT-HEALTH

MH-057 - project-world-model-project-health. Execute project health as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-059 PROJECT-WORLD-MODEL-CROSS-PROJECT-IMPACT

MH-059 - project-world-model-cross-project-impact. Execute cross project impact as an auditable capability within the Project World Model layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

PROJECT_WORLD_MODEL

Trigger	Name the condition that activates PROJECT_WORLD_MODEL; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Project World Model - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Repository Code - Skills 1 to 8

CATEGORY ROLE

Repository Code contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-061 REPOSITORY-CODE-DOMAIN-MIGRATION

MH-061 - repository-code-domain-migration. Execute domain migration as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-063 REPOSITORY-CODE-SYMBOL-TRACING

MH-063 - repository-code-symbol-tracing. Execute symbol tracing as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-065 REPOSITORY-CODE-CONFIG-DRIFT

MH-065 - repository-code-config-drift. Execute config drift as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-067 REPOSITORY-CODE-API-CONTRACT-MIGRATION

MH-067 - repository-code-api-contract-migration. Execute api contract migration as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-062 REPOSITORY-CODE-STRUCTURE-PROFILING

MH-062 - repository-code-structure-profiling. Execute structure profiling as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-064 REPOSITORY-CODE-SAFE-REFACTORING

MH-064 - repository-code-safe-refactoring. Execute safe refactoring as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-066 REPOSITORY-CODE-DEAD-CODE-AUDIT

MH-066 - repository-code-dead-code-audit. Execute dead code audit as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-068 REPOSITORY-CODE-SCHEMA-CHANGE-PLANNING

MH-068 - repository-code-schema-change-planning. Execute schema change planning as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

REPOSITORY_CODE

Trigger	Name the condition that activates REPOSITORY_CODE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Repository Code - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Repository Code - Skills 9 to 15

CATEGORY ROLE

Repository Code contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-069 REPOSITORY-CODE-DEPENDENCY-UPGRADE

MH-069 - repository-code-dependency-upgrade. Execute dependency upgrade as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-071 REPOSITORY-CODE-GENERATED-FILE-EXCLUSION

MH-071 - repository-code-generated-file-exclusion. Execute generated file exclusion as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-073 REPOSITORY-CODE-MERGE-CONFLICT-REPAIR

MH-073 - repository-code-merge-conflict-repair. Execute merge conflict repair as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-075 REPOSITORY-CODE-RELEASE-READINESS

MH-075 - repository-code-release-readiness. Execute release readiness as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-070 REPOSITORY-CODE-SECRET-LEAK-SCAN

MH-070 - repository-code-secret-leak-scan. Execute secret leak scan as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-072 REPOSITORY-CODE-BRANCH-STRATEGY

MH-072 - repository-code-branch-strategy. Execute branch strategy as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-074 REPOSITORY-CODE-REGRESSION-SURFACE

MH-074 - repository-code-regression-surface. Execute regression surface as an auditable capability within the Repository and Code Intelligence layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

REPOSITORY_CODE

Trigger	Name the condition that activates REPOSITORY_CODE; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Repository Code - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Knowledge Memory - Skills 1 to 8

CATEGORY ROLE

Knowledge Memory contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-076 KNOWLEDGE-MEMORY-HYBRID-RETRIEVAL

MH-076 - knowledge-memory-hybrid-retrieval. Execute hybrid retrieval as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-078 KNOWLEDGE-MEMORY-KEYWORD-QUERY-COMPILATION

MH-078 - knowledge-memory-keyword-query-compilation. Execute keyword query compilation as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-080 KNOWLEDGE-MEMORY-CONTEXT-BUDGETING

MH-080 - knowledge-memory-context-budgeting. Execute context budgeting as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-082 KNOWLEDGE-MEMORY-MEMORY-WRITE-DECISION

MH-082 - knowledge-memory-memory-write-decision. Execute memory write decision as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-077 KNOWLEDGE-MEMORY-SEMANTIC-QUERY-EXPANSION

MH-077 - knowledge-memory-semantic-query-expansion. Execute semantic query expansion as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-079 KNOWLEDGE-MEMORY-KNOWLEDGE-GRAPH-TRAVERSAL

MH-079 - knowledge-memory-knowledge-graph-traversal. Execute knowledge graph traversal as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-081 KNOWLEDGE-MEMORY-DOCUMENT-CHUNKING

MH-081 - knowledge-memory-document-chunking. Execute document chunking as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-083 KNOWLEDGE-MEMORY-MEMORY-FORGET-DECISION

MH-083 - knowledge-memory-memory-forget-decision. Execute memory forget decision as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

KNOWLEDGE_MEMORY

Trigger	Name the condition that activates KNOWLEDGE_MEMORY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Knowledge Memory - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Knowledge Memory - Skills 9 to 15

CATEGORY ROLE

Knowledge Memory contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-084 KNOWLEDGE-MEMORY-EPISODIC-SUMMARIZATION

MH-084 - knowledge-memory-episodic-summarization. Execute episodic summarization as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-086 KNOWLEDGE-MEMORY-FRESHNESS-CHECKING

MH-086 - knowledge-memory-freshness-checking. Execute freshness checking as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-088 KNOWLEDGE-MEMORY-DUPLICATE-MERGING

MH-088 - knowledge-memory-duplicate-merging. Execute duplicate merging as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-090 KNOWLEDGE-MEMORY-RETRIEVAL-EVALUATION

MH-090 - knowledge-memory-retrieval-evaluation. Execute retrieval evaluation as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-085 KNOWLEDGE-MEMORY-SEMANTIC-CONSOLIDATION

MH-085 - knowledge-memory-semantic-consolidation. Execute semantic consolidation as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-087 KNOWLEDGE-MEMORY-PROVENANCE-PRESERVATION

MH-087 - knowledge-memory-provenance-preservation. Execute provenance preservation as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-089 KNOWLEDGE-MEMORY-MEMORY-CONFLICT-RESOLUTION

MH-089 - knowledge-memory-memory-conflict-resolution. Execute memory conflict resolution as an auditable capability within the Knowledge Retrieval and Memory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

KNOWLEDGE_MEMORY

Trigger	Name the condition that activates KNOWLEDGE_MEMORY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Knowledge Memory - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Cognitive Routing - Skills 1 to 8

CATEGORY ROLE

Cognitive Routing contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-091 COGNITIVE-ROUTING-SKILL-RELEVANCE

MH-091 - cognitive-routing-skill-relevance. Execute skill relevance as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-093 COGNITIVE-ROUTING-COMPATIBILITY-CHECKING

MH-093 - cognitive-routing-compatibility-checking. Execute compatibility checking as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-095 COGNITIVE-ROUTING-HISTORICAL-SUCCESS

MH-095 - cognitive-routing-historical-success. Execute historical success as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-097 COGNITIVE-ROUTING-SECURITY-RISK

MH-097 - cognitive-routing-security-risk. Execute security risk as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-092 COGNITIVE-ROUTING-UTILITY-ESTIMATION

MH-092 - cognitive-routing-utility-estimation. Execute utility estimation as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-094 COGNITIVE-ROUTING-EVIDENCE-AVAILABILITY

MH-094 - cognitive-routing-evidence-availability. Execute evidence availability as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-096 COGNITIVE-ROUTING-EXECUTION-COST

MH-096 - cognitive-routing-execution-cost. Execute execution cost as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-098 COGNITIVE-ROUTING-CONTEXT-COST

MH-098 - cognitive-routing-context-cost. Execute context cost as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

COGNITIVE_ROUTING

Trigger	Name the condition that activates COGNITIVE_ROUTING; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Cognitive Routing - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Cognitive Routing - Skills 9 to 15

CATEGORY ROLE

Cognitive Routing contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-099 COGNITIVE-ROUTING-REDUNDANCY-PENALTY

MH-099 - cognitive-routing-redundancy-penalty. Execute redundancy penalty as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-101 COGNITIVE-ROUTING-TASK-COMPLEXITY

MH-101 - cognitive-routing-task-complexity. Execute task complexity as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-103 COGNITIVE-ROUTING-CRITICAL-PATH

MH-103 - cognitive-routing-critical-path. Execute critical path as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-105 COGNITIVE-ROUTING-PLAN-STABILITY

MH-105 - cognitive-routing-plan-stability. Execute plan stability as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-100 COGNITIVE-ROUTING-MODEL-ROUTING

MH-100 - cognitive-routing-model-routing. Execute model routing as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-102 COGNITIVE-ROUTING-MISSION-DECOMPOSITION

MH-102 - cognitive-routing-mission-decomposition. Execute mission decomposition as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-104 COGNITIVE-ROUTING-PARALLEL-PLANNING

MH-104 - cognitive-routing-parallel-planning. Execute parallel planning as an auditable capability within the Cognitive Routing and Planning layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

COGNITIVE_ROUTING

Trigger	Name the condition that activates COGNITIVE_ROUTING; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Cognitive Routing - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Autonomous Execution - Skills 1 to 8

CATEGORY ROLE

Autonomous Execution contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-106 AUTONOMOUS-EXECUTION-EXECUTION-STATE-MACHINE

MH-106 - autonomous-execution-execution-state-machine. Execute execution state machine as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-108 AUTONOMOUS-EXECUTION-PREFLIGHT-CHECKS

MH-108 - autonomous-execution-preflight-checks. Execute preflight checks as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-110 AUTONOMOUS-EXECUTION-TRANSACTIONAL-EDITING

MH-110 - autonomous-execution-transactional-editing. Execute transactional editing as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-112 AUTONOMOUS-EXECUTION-MCP-ORCHESTRATION

MH-112 - autonomous-execution-mcp-orchestration. Execute mcp orchestration as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-107 AUTONOMOUS-EXECUTION-TOOL-ARGUMENT-BUILDING

MH-107 - autonomous-execution-tool-argument-building. Execute tool argument building as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-109 AUTONOMOUS-EXECUTION-CHECKPOINT-MANAGEMENT

MH-109 - autonomous-execution-checkpoint-management. Execute checkpoint management as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-111 AUTONOMOUS-EXECUTION-BROWSER-WORKFLOW

MH-111 - autonomous-execution-browser-workflow. Execute browser workflow as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-113 AUTONOMOUS-EXECUTION-PARALLEL-EXECUTION

MH-113 - autonomous-execution-parallel-execution. Execute parallel execution as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

AUTONOMOUS_EXECUTION

Trigger	Name the condition that activates AUTONOMOUS_EXECUTION; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Autonomous Execution - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Autonomous Execution - Skills 9 to 15

CATEGORY ROLE

Autonomous Execution contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-114 AUTONOMOUS-EXECUTION-JOB-SUPERVISION

MH-114 - autonomous-execution-job-supervision. Execute job supervision as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-116 AUTONOMOUS-EXECUTION-ENVIRONMENT-COORDINATION

MH-116 - autonomous-execution-environment-coordination. Execute environment coordination as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-118 AUTONOMOUS-EXECUTION-IDEMPOTENCY

MH-118 - autonomous-execution-idempotency. Execute idempotency as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-120 AUTONOMOUS-EXECUTION-COMPLETION-EVIDENCE

MH-120 - autonomous-execution-completion-evidence. Execute completion evidence as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-115 AUTONOMOUS-EXECUTION-ARTIFACT-MANAGEMENT

MH-115 - autonomous-execution-artifact-management. Execute artifact management as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-117 AUTONOMOUS-EXECUTION-SANDBOX-CONTROL

MH-117 - autonomous-execution-sandbox-control. Execute sandbox control as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-119 AUTONOMOUS-EXECUTION-TIMEOUT-MANAGEMENT

MH-119 - autonomous-execution-timeout-management. Execute timeout management as an auditable capability within the Autonomous Execution layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

AUTONOMOUS_EXECUTION

Trigger	Name the condition that activates AUTONOMOUS_EXECUTION; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Autonomous Execution - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Testing Verification - Skills 1 to 8

CATEGORY ROLE

Testing Verification contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-121 TESTING-VERIFICATION-ACCEPTANCE-CRITERIA

MH-121 - testing-verification-acceptance-criteria. Execute acceptance criteria as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-123 TESTING-VERIFICATION-INTEGRATION-TESTING

MH-123 - testing-verification-integration-testing. Execute integration testing as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-125 TESTING-VERIFICATION-PROPERTY-TESTING

MH-125 - testing-verification-property-testing. Execute property testing as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-127 TESTING-VERIFICATION-TYPE-SAFETY

MH-127 - testing-verification-type-safety. Execute type safety as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-122 TESTING-VERIFICATION-UNIT-TESTING

MH-122 - testing-verification-unit-testing. Execute unit testing as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-124 TESTING-VERIFICATION-END-TO-END-TESTING

MH-124 - testing-verification-end-to-end-testing. Execute end to end testing as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-126 TESTING-VERIFICATION-STATIC-ANALYSIS

MH-126 - testing-verification-static-analysis. Execute static analysis as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-128 TESTING-VERIFICATION-VISUAL-REGRESSION

MH-128 - testing-verification-visual-regression. Execute visual regression as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

TESTING_VERIFICATION

Trigger	Name the condition that activates TESTING_VERIFICATION; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Testing Verification - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Testing Verification - Skills 9 to 15

CATEGORY ROLE

Testing Verification contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-129 TESTING-VERIFICATION-DEPLOYMENT-SMOKE-TEST

MH-129 - testing-verification-deployment-smoke-test. Execute deployment smoke test as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-131 TESTING-VERIFICATION-SECURITY-TESTING

MH-131 - testing-verification-security-testing. Execute security testing as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-133 TESTING-VERIFICATION-ARTIFACT-DIFF

MH-133 - testing-verification-artifact-diff. Execute artifact diff as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-135 TESTING-VERIFICATION-VERIFICATION-REPORT

MH-135 - testing-verification-verification-report. Execute verification report as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-130 TESTING-VERIFICATION-DATA-INTEGRITY

MH-130 - testing-verification-data-integrity. Execute data integrity as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-132 TESTING-VERIFICATION-PERFORMANCE-BENCHMARK

MH-132 - testing-verification-performance-benchmark. Execute performance benchmark as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-134 TESTING-VERIFICATION-COMPLETION-VERIFICATION

MH-134 - testing-verification-completion-verification. Execute completion verification as an auditable capability within the Testing and Verification layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

TESTING_VERIFICATION

Trigger	Name the condition that activates TESTING_VERIFICATION; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Testing Verification - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Failure Recovery - Skills 1 to 8

CATEGORY ROLE

Failure Recovery contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-136 FAILURE-RECOVERY-FAILURE-CLASSIFICATION

MH-136 - failure-recovery-failure-classification. Execute failure classification as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-138 FAILURE-RECOVERY-MINIMAL-REPRODUCTION

MH-138 - failure-recovery-minimal-reproduction. Execute minimal reproduction as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-140 FAILURE-RECOVERY-FALLBACK-TOOL

MH-140 - failure-recovery-fallback-tool. Execute fallback tool as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-142 FAILURE-RECOVERY-PARTIAL-PROGRESS-RECOVERY

MH-142 - failure-recovery-partial-progress-recovery. Execute partial progress recovery as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-137 FAILURE-RECOVERY-ROOT-CAUSE-HYPOTHESIS

MH-137 - failure-recovery-root-cause-hypothesis. Execute root cause hypothesis as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-139 FAILURE-RECOVERY-RETRY-POLICY

MH-139 - failure-recovery-retry-policy. Execute retry policy as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-141 FAILURE-RECOVERY-ROLLBACK-PLAN

MH-141 - failure-recovery-rollback-plan. Execute rollback plan as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-143 FAILURE-RECOVERY-CORRUPTED-STATE-DETECTION

MH-143 - failure-recovery-corrupted-state-detection. Execute corrupted state detection as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

FAILURE_RECOVERY

Trigger	Name the condition that activates FAILURE_RECOVERY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Failure Recovery - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Failure Recovery - Skills 9 to 15

CATEGORY ROLE

Failure Recovery contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-144 FAILURE-RECOVERY-DEPENDENCY-FAILURE

MH-144 - failure-recovery-dependency-failure. Execute dependency failure as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-146 FAILURE-RECOVERY-PERMISSION-RECOVERY

MH-146 - failure-recovery-permission-recovery. Execute permission recovery as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-148 FAILURE-RECOVERY-LOOP-BREAKING

MH-148 - failure-recovery-loop-breaking. Execute loop breaking as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-150 FAILURE-RECOVERY-FAILURE-TO-SKILL

MH-150 - failure-recovery-failure-to-skill. Execute failure to skill as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-145 FAILURE-RECOVERY-NETWORK-RECOVERY

MH-145 - failure-recovery-network-recovery. Execute network recovery as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-147 FAILURE-RECOVERY-RESOURCE-EXHAUSTION

MH-147 - failure-recovery-resource-exhaustion. Execute resource exhaustion as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-149 FAILURE-RECOVERY-INCIDENT-POSTMORTEM

MH-149 - failure-recovery-incident-postmortem. Execute incident postmortem as an auditable capability within the Failure Recovery layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

FAILURE_RECOVERY

Trigger	Name the condition that activates FAILURE_RECOVERY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Failure Recovery - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Security Authorization - Skills 1 to 8

CATEGORY ROLE

Security Authorization contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-151 SECURITY-AUTHORIZATION-LEAST-PRIVILEGE

MH-151 - security-authorization-least-privilege. Execute least privilege as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-153 SECURITY-AUTHORIZATION-CREDENTIAL-GUARDIAN

MH-153 - security-authorization-credential-guardian. Execute credential guardian as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-155 SECURITY-AUTHORIZATION-DATA-CLASSIFICATION

MH-155 - security-authorization-data-classification. Execute data classification as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-157 SECURITY-AUTHORIZATION-EXTERNAL-SIDE-EFFECT-GATE

MH-157 - security-authorization-external-side-effect-gate. Execute external side effect gate as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-152 SECURITY-AUTHORIZATION-DESTRUCTIVE-ACTION-GATE

MH-152 - security-authorization-destructive-action-gate. Execute destructive action gate as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-154 SECURITY-AUTHORIZATION-SECRET-REDACTION

MH-154 - security-authorization-secret-redaction. Execute secret redaction as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-156 SECURITY-AUTHORIZATION-PRIVACY-BOUNDARIES

MH-156 - security-authorization-privacy-boundaries. Execute privacy boundaries as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-158 SECURITY-AUTHORIZATION-FINANCIAL-APPROVAL

MH-158 - security-authorization-financial-approval. Execute financial approval as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

SECURITY_AUTHORIZATION

Trigger	Name the condition that activates SECURITY_AUTHORIZATION; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Security Authorization - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Security Authorization - Skills 9 to 15

CATEGORY ROLE

Security Authorization contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-159 SECURITY-AUTHORIZATION-PRODUCTION-APPROVAL

MH-159 - security-authorization-production-approval. Execute production approval as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-161 SECURITY-AUTHORIZATION-TOOL-TRUST-EVALUATION

MH-161 - security-authorization-tool-trust-evaluation. Execute tool trust evaluation as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-163 SECURITY-AUTHORIZATION-SANDBOX-POLICY

MH-163 - security-authorization-sandbox-policy. Execute sandbox policy as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-165 SECURITY-AUTHORIZATION-INCIDENT-ESCALATION

MH-165 - security-authorization-incident-escalation. Execute incident escalation as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-160 SECURITY-AUTHORIZATION-PROMPT-INJECTION-DETECTION

MH-160 - security-authorization-prompt-injection-detection. Execute prompt injection detection as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-162 SECURITY-AUTHORIZATION-SUPPLY-CHAIN-RISK

MH-162 - security-authorization-supply-chain-risk. Execute supply chain risk as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-164 SECURITY-AUTHORIZATION-AUDIT-TRAIL

MH-164 - security-authorization-audit-trail. Execute audit trail as an auditable capability within the Security and Authorization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

SECURITY_AUTHORIZATION

Trigger	Name the condition that activates SECURITY_AUTHORIZATION; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Security Authorization - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Infrastructure Gpu - Skills 1 to 8

CATEGORY ROLE

Infrastructure Gpu contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-166 INFRASTRUCTURE-GPU-GPU-PROFILING

MH-166 - infrastructure-gpu-gpu-profiling. Execute gpu profiling as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-168 INFRASTRUCTURE-GPU-QUANTIZATION-SELECTION

MH-168 - infrastructure-gpu-quantization-selection. Execute quantization selection as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-170 INFRASTRUCTURE-GPU-BATCHING-TUNING

MH-170 - infrastructure-gpu-batching-tuning. Execute batching tuning as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-172 INFRASTRUCTURE-GPU-PREFIX-CACHE

MH-172 - infrastructure-gpu-prefix-cache. Execute prefix cache as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-167 INFRASTRUCTURE-GPU-VRAM-FIT

MH-167 - infrastructure-gpu-vram-fit. Execute vram fit as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-169 INFRASTRUCTURE-GPU-VLLM-CONFIGURATION

MH-169 - infrastructure-gpu-vllm-configuration. Execute vllm configuration as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-171 INFRASTRUCTURE-GPU-KV-CACHE-PLANNING

MH-171 - infrastructure-gpu-kv-cache-planning. Execute kv cache planning as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-173 INFRASTRUCTURE-GPU-TENSOR-PARALLELISM

MH-173 - infrastructure-gpu-tensor-parallelism. Execute tensor parallelism as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

INFRASTRUCTURE_GPU

Trigger	Name the condition that activates INFRASTRUCTURE_GPU; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Infrastructure Gpu - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Infrastructure Gpu - Skills 9 to 15

CATEGORY ROLE

Infrastructure Gpu contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-174 INFRASTRUCTURE-GPU-OOM-RECOVERY

MH-174 - infrastructure-gpu-oom-recovery. Execute oom recovery as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-176 INFRASTRUCTURE-GPU-INFERENCE-QUEUE

MH-176 - infrastructure-gpu-inference-queue. Execute inference queue as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-178 INFRASTRUCTURE-GPU-ENDPOINT-HEALTH

MH-178 - infrastructure-gpu-endpoint-health. Execute endpoint health as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-180 INFRASTRUCTURE-GPU-INFERENCE-RESILIENCE

MH-180 - infrastructure-gpu-inference-resilience. Execute inference resilience as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-175 INFRASTRUCTURE-GPU-THROUGHPUT-LATENCY

MH-175 - infrastructure-gpu-throughput-latency. Execute throughput latency as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-177 INFRASTRUCTURE-GPU-GPU-OBSERVABILITY

MH-177 - infrastructure-gpu-gpu-observability. Execute gpu observability as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-179 INFRASTRUCTURE-GPU-MODEL-HOT-SWAP

MH-179 - infrastructure-gpu-model-hot-swap. Execute model hot swap as an auditable capability within the Infrastructure and GPU Optimization layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

INFRASTRUCTURE_GPU

Trigger	Name the condition that activates INFRASTRUCTURE_GPU; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Infrastructure Gpu - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Data Training - Skills 1 to 8

CATEGORY ROLE

Data Training contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-181 DATA-TRAINING-TRACE-EXTRACTION

MH-181 - data-training-trace-extraction. Execute trace extraction as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-183 DATA-TRAINING-TOOL-CALL-DATASET

MH-183 - data-training-tool-call-dataset. Execute tool call dataset as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-185 DATA-TRAINING-PREFERENCE-PAIRS

MH-185 - data-training-preference-pairs. Execute preference pairs as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-187 DATA-TRAINING-DATA-LEAKAGE

MH-187 - data-training-data-leakage. Execute data leakage as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-182 DATA-TRAINING-EXAMPLE-NORMALIZATION

MH-182 - data-training-example-normalization. Execute example normalization as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-184 DATA-TRAINING-SKILL-ROUTING-DATASET

MH-184 - data-training-skill-routing-dataset. Execute skill routing dataset as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-186 DATA-TRAINING-HARD-NEGATIVES

MH-186 - data-training-hard-negatives. Execute hard negatives as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-188 DATA-TRAINING-DEDUPLICATION

MH-188 - data-training-deduplication. Execute deduplication as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

DATA_TRAINING

Trigger	Name the condition that activates DATA_TRAINING; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Data Training - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Data Training - Skills 9 to 15

CATEGORY ROLE

Data Training contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-189 DATA-TRAINING-CURRICULUM-PLANNING

MH-189 - data-training-curriculum-planning. Execute curriculum planning as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-191 DATA-TRAINING-ADAPTER-MERGING

MH-191 - data-training-adapter-merging. Execute adapter merging as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-193 DATA-TRAINING-MODEL-EVALUATION

MH-193 - data-training-model-evaluation. Execute model evaluation as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-195 DATA-TRAINING-TRAINING-POSTMORTEM

MH-195 - data-training-training-postmortem. Execute training postmortem as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-190 DATA-TRAINING-QLORA-CONFIGURATION

MH-190 - data-training-qlora-configuration. Execute qlora configuration as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-192 DATA-TRAINING-FORGETTING-TESTS

MH-192 - data-training-forgetting-tests. Execute forgetting tests as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-194 DATA-TRAINING-CHECKPOINT-PROMOTION

MH-194 - data-training-checkpoint-promotion. Execute checkpoint promotion as an auditable capability within the Data and Model Training layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

DATA_TRAINING

Trigger	Name the condition that activates DATA_TRAINING; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Data Training - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Web Deployment - Skills 1 to 8

CATEGORY ROLE

Web Deployment contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-196 WEB-DEPLOYMENT-WEB-BOOTSTRAP

MH-196 - web-deployment-web-bootstrap. Execute web bootstrap as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-198 WEB-DEPLOYMENT-BACKEND-PLANNING

MH-198 - web-deployment-backend-planning. Execute backend planning as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-200 WEB-DEPLOYMENT-DOMAIN-MIGRATION

MH-200 - web-deployment-domain-migration. Execute domain migration as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-202 WEB-DEPLOYMENT-CI-CD

MH-202 - web-deployment-ci-cd. Execute ci cd as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-197 WEB-DEPLOYMENT-FRONTEND-ARCHITECTURE

MH-197 - web-deployment-frontend-architecture. Execute frontend architecture as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-199 WEB-DEPLOYMENT-ENVIRONMENT-CONFIG

MH-199 - web-deployment-environment-config. Execute environment config as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-201 WEB-DEPLOYMENT-DNS-VALIDATION

MH-201 - web-deployment-dns-validation. Execute dns validation as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-203 WEB-DEPLOYMENT-PREVIEW-DEPLOYMENT

MH-203 - web-deployment-preview-deployment. Execute preview deployment as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

WEB_DEPLOYMENT

Trigger	Name the condition that activates WEB_DEPLOYMENT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Web Deployment - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Web Deployment - Skills 9 to 15

CATEGORY ROLE

Web Deployment contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-204 WEB-DEPLOYMENT-PRODUCTION-DEPLOYMENT

MH-204 - web-deployment-production-deployment. Execute production deployment as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-206 WEB-DEPLOYMENT-OBSERVABILITY-INSTRUMENTATION

MH-206 - web-deployment-observability-instrumentation. Execute observability instrumentation as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-208 WEB-DEPLOYMENT-TECHNICAL-SEO

MH-208 - web-deployment-technical-seo. Execute technical seo as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-210 WEB-DEPLOYMENT-WEB-PERFORMANCE

MH-210 - web-deployment-web-performance. Execute web performance as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-205 WEB-DEPLOYMENT-DEPLOYMENT-ROLLBACK

MH-205 - web-deployment-deployment-rollback. Execute deployment rollback as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-207 WEB-DEPLOYMENT-ANALYTICS-EVENTS

MH-207 - web-deployment-analytics-events. Execute analytics events as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-209 WEB-DEPLOYMENT-ACCESSIBILITY-AUDIT

MH-209 - web-deployment-accessibility-audit. Execute accessibility audit as an auditable capability within the Web Product and Deployment layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

WEB_DEPLOYMENT

Trigger	Name the condition that activates WEB_DEPLOYMENT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Web Deployment - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Business Growth - Skills 1 to 8

CATEGORY ROLE

Business Growth contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-211 BUSINESS-GROWTH-CUSTOMER-PROBLEM

MH-211 - business-growth-customer-problem. Execute customer problem as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-213 BUSINESS-GROWTH-OFFER-ARCHITECTURE

MH-213 - business-growth-offer-architecture. Execute offer architecture as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-215 BUSINESS-GROWTH-REVENUE-HYPOTHESIS

MH-215 - business-growth-revenue-hypothesis. Execute revenue hypothesis as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-217 BUSINESS-GROWTH-LEAD-PRIORITIZATION

MH-217 - business-growth-lead-prioritization. Execute lead prioritization as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-212 BUSINESS-GROWTH-VALUE-PROPOSITION

MH-212 - business-growth-value-proposition. Execute value proposition as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-214 BUSINESS-GROWTH-BUSINESS-MODEL

MH-214 - business-growth-business-model. Execute business model as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-216 BUSINESS-GROWTH-PRICING-EXPERIMENT

MH-216 - business-growth-pricing-experiment. Execute pricing experiment as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-218 BUSINESS-GROWTH-FUNNEL-DIAGNOSIS

MH-218 - business-growth-funnel-diagnosis. Execute funnel diagnosis as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

BUSINESS_GROWTH

Trigger	Name the condition that activates BUSINESS_GROWTH; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Business Growth - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Business Growth - Skills 9 to 15

CATEGORY ROLE

Business Growth contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-219 BUSINESS-GROWTH-GROWTH-LOOP

MH-219 - business-growth-growth-loop. Execute growth loop as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-221 BUSINESS-GROWTH-UNIT-ECONOMICS

MH-221 - business-growth-unit-economics. Execute unit economics as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-223 BUSINESS-GROWTH-COMPETITIVE-POSITIONING

MH-223 - business-growth-competitive-positioning. Execute competitive positioning as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-225 BUSINESS-GROWTH-PROJECT-PRIORITIZATION

MH-225 - business-growth-project-prioritization. Execute project prioritization as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-220 BUSINESS-GROWTH-RETENTION-MECHANISM

MH-220 - business-growth-retention-mechanism. Execute retention mechanism as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-222 BUSINESS-GROWTH-MARKET-SIGNALS

MH-222 - business-growth-market-signals. Execute market signals as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-224 BUSINESS-GROWTH-BUSINESS-IMPACT

MH-224 - business-growth-business-impact. Execute business impact as an auditable capability within the Business Value and Growth layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

BUSINESS_GROWTH

Trigger	Name the condition that activates BUSINESS_GROWTH; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Business Growth - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Research Theory - Skills 1 to 8

CATEGORY ROLE

Research Theory contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-226 RESEARCH-THEORY-QUESTION-REFINEMENT

MH-226 - research-theory-question-refinement. Execute question refinement as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-228 RESEARCH-THEORY-CONSTRUCT-DEFINITION

MH-228 - research-theory-construct-definition. Execute construct definition as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-230 RESEARCH-THEORY-FORMALISM-SELECTION

MH-230 - research-theory-formalism-selection. Execute formalism selection as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-232 RESEARCH-THEORY-EXPERIMENTAL-DESIGN

MH-232 - research-theory-experimental-design. Execute experimental design as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-227 RESEARCH-THEORY-LITERATURE-MAP

MH-227 - research-theory-literature-map. Execute literature map as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-229 RESEARCH-THEORY-MECHANISM-MODEL

MH-229 - research-theory-mechanism-model. Execute mechanism model as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-231 RESEARCH-THEORY-FALSIFIABILITY-AUDIT

MH-231 - research-theory-falsifiability-audit. Execute falsifiability audit as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-233 RESEARCH-THEORY-MEASUREMENT-INSTRUMENT

MH-233 - research-theory-measurement-instrument. Execute measurement instrument as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

RESEARCH_THEORY

Trigger	Name the condition that activates RESEARCH_THEORY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Research Theory - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Research Theory - Skills 9 to 15

CATEGORY ROLE

Research Theory contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-234 RESEARCH-THEORY-ARGUMENT-STRUCTURE

MH-234 - research-theory-argument-structure. Execute argument structure as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-236 RESEARCH-THEORY-NOVELTY-CALIBRATION

MH-236 - research-theory-novelty-calibration. Execute novelty calibration as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-238 RESEARCH-THEORY-RESEARCH-GAP

MH-238 - research-theory-research-gap. Execute research gap as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-240 RESEARCH-THEORY-PROGRAM-ROADMAP

MH-240 - research-theory-program-roadmap. Execute program roadmap as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-235 RESEARCH-THEORY-CITATION-INTEGRITY

MH-235 - research-theory-citation-integrity. Execute citation integrity as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-237 RESEARCH-THEORY-THEORY-COMPARISON

MH-237 - research-theory-theory-comparison. Execute theory comparison as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-239 RESEARCH-THEORY-PAPER-ARCHITECTURE

MH-239 - research-theory-paper-architecture. Execute paper architecture as an auditable capability within the Research and Theory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

RESEARCH_THEORY

Trigger	Name the condition that activates RESEARCH_THEORY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Research Theory - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Content Brand - Skills 1 to 8

CATEGORY ROLE

Content Brand contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-241 CONTENT-BRAND-BRAND-CANONICALIZATION

MH-241 - content-brand-brand-canonicalization. Execute brand canonicalization as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-243 CONTENT-BRAND-JARGON-REDUCTION

MH-243 - content-brand-jargon-reduction. Execute jargon reduction as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-245 CONTENT-BRAND-CONTENT-PILLARS

MH-245 - content-brand-content-pillars. Execute content pillars as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-247 CONTENT-BRAND-ARTICLE-TO-SOCIAL

MH-247 - content-brand-article-to-social. Execute article to social as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-242 CONTENT-BRAND-AUDIENCE-LANGUAGE

MH-242 - content-brand-audience-language. Execute audience language as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-244 CONTENT-BRAND-AUTHORITY-NARRATIVE

MH-244 - content-brand-authority-narrative. Execute authority narrative as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-246 CONTENT-BRAND-RESEARCH-TO-ARTICLE

MH-246 - content-brand-research-to-article. Execute research to article as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-248 CONTENT-BRAND-CHANNEL-CONSISTENCY

MH-248 - content-brand-channel-consistency. Execute channel consistency as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

CONTENT_BRAND

Trigger	Name the condition that activates CONTENT_BRAND; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Content Brand - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Content Brand - Skills 9 to 15

CATEGORY ROLE

Content Brand contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-249 CONTENT-BRAND-EDITORIAL-CALENDAR

MH-249 - content-brand-editorial-calendar. Execute editorial calendar as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-251 CONTENT-BRAND-EVIDENCE-CHECKING

MH-251 - content-brand-evidence-checking. Execute evidence checking as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-253 CONTENT-BRAND-VISUAL-BRIEF

MH-253 - content-brand-visual-brief. Execute visual brief as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-255 CONTENT-BRAND-PERFORMANCE-LEARNING

MH-255 - content-brand-performance-learning. Execute performance learning as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-250 CONTENT-BRAND-CONVERSION-COPY

MH-250 - content-brand-conversion-copy. Execute conversion copy as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-252 CONTENT-BRAND-CONTENT-REPURPOSING

MH-252 - content-brand-content-repurposing. Execute content repurposing as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-254 CONTENT-BRAND-PUBLISHING-GATE

MH-254 - content-brand-publishing-gate. Execute publishing gate as an auditable capability within the Content and Brand layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

CONTENT_BRAND

Trigger	Name the condition that activates CONTENT_BRAND; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Content Brand - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Multi Agent - Skills 1 to 8

CATEGORY ROLE

Multi Agent contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-256 MULTI-AGENT-ROLE-DESIGN

MH-256 - multi-agent-role-design. Execute role design as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-258 MULTI-AGENT-TASK-DELEGATION

MH-258 - multi-agent-task-delegation. Execute task delegation as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-260 MULTI-AGENT-DEPENDENCY-COORDINATION

MH-260 - multi-agent-dependency-coordination. Execute dependency coordination as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-262 MULTI-AGENT-RESULT-MERGING

MH-262 - multi-agent-result-merging. Execute result merging as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-257 MULTI-AGENT-CAPABILITY-MATCHING

MH-257 - multi-agent-capability-matching. Execute capability matching as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-259 MULTI-AGENT-SHARED-CONTEXT

MH-259 - multi-agent-shared-context. Execute shared context as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-261 MULTI-AGENT-PARALLEL-SCHEDULING

MH-261 - multi-agent-parallel-scheduling. Execute parallel scheduling as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-263 MULTI-AGENT-CONFLICT-ARBITRATION

MH-263 - multi-agent-conflict-arbitration. Execute conflict arbitration as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

MULTI_AGENT

Trigger	Name the condition that activates MULTI_AGENT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Multi Agent - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Multi Agent - Skills 9 to 15

CATEGORY ROLE

Multi Agent contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-264 MULTI-AGENT-CONSENSUS-BUILDING

MH-264 - multi-agent-consensus-building. Execute consensus building as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-266 MULTI-AGENT-VERIFIER-AGENT

MH-266 - multi-agent-verifier-agent. Execute verifier agent as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-268 MULTI-AGENT-LOOP-TERMINATION

MH-268 - multi-agent-loop-termination. Execute loop termination as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-270 MULTI-AGENT-TEAM-PERFORMANCE

MH-270 - multi-agent-team-performance. Execute team performance as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-265 MULTI-AGENT-CRITIC-AGENT

MH-265 - multi-agent-critic-agent. Execute critic agent as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-267 MULTI-AGENT-SPECIALIST-AGENT

MH-267 - multi-agent-specialist-agent. Execute specialist agent as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-269 MULTI-AGENT-BUDGET-ALLOCATION

MH-269 - multi-agent-budget-allocation. Execute budget allocation as an auditable capability within the Multi-Agent Coordination layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

MULTI_AGENT

Trigger	Name the condition that activates MULTI_AGENT; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Multi Agent - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Observability - Skills 1 to 8

CATEGORY ROLE

Observability contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-271 OBSERVABILITY-TRACE-INSTRUMENTATION

MH-271 - observability-trace-instrumentation. Execute trace instrumentation as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-273 OBSERVABILITY-TOOL-CALL-ACCURACY

MH-273 - observability-tool-call-accuracy. Execute tool call accuracy as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-275 OBSERVABILITY-MISSION-SUCCESS

MH-275 - observability-mission-success. Execute mission success as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-277 OBSERVABILITY-TOKEN-COST

MH-277 - observability-token-cost. Execute token cost as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-272 OBSERVABILITY-SELECTION-METRICS

MH-272 - observability-selection-metrics. Execute selection metrics as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-274 OBSERVABILITY-RETRIEVAL-PRECISION

MH-274 - observability-retrieval-precision. Execute retrieval precision as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-276 OBSERVABILITY-LATENCY-ANALYSIS

MH-276 - observability-latency-analysis. Execute latency analysis as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-278 OBSERVABILITY-GPU-EFFICIENCY

MH-278 - observability-gpu-efficiency. Execute gpu efficiency as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

OBSERVABILITY

Trigger	Name the condition that activates OBSERVABILITY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Observability - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Observability - Skills 9 to 15

CATEGORY ROLE

Observability contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-279 OBSERVABILITY-HALLUCINATION-ESTIMATION

MH-279 - observability-hallucination-estimation. Execute hallucination estimation as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-281 OBSERVABILITY-FAILURE-RECURRENCE

MH-281 - observability-failure-recurrence. Execute failure recurrence as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-283 OBSERVABILITY-BENCHMARK-MANAGEMENT

MH-283 - observability-benchmark-management. Execute benchmark management as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-285 OBSERVABILITY-OPERATIONAL-ALERTS

MH-285 - observability-operational-alerts. Execute operational alerts as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-280 OBSERVABILITY-VERIFICATION-COVERAGE

MH-280 - observability-verification-coverage. Execute verification coverage as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-282 OBSERVABILITY-QUALITY-REGRESSION

MH-282 - observability-quality-regression. Execute quality regression as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-284 OBSERVABILITY-EVALUATION-DASHBOARD

MH-284 - observability-evaluation-dashboard. Execute evaluation dashboard as an auditable capability within the Observability and Evaluation layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=medium. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

OBSERVABILITY

Trigger	Name the condition that activates OBSERVABILITY; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Observability - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Evolution Lab - Skills 1 to 8

CATEGORY ROLE

Evolution Lab contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-286 EVOLUTION-LAB-HARD-TASK-GENERATION

MH-286 - evolution-lab-hard-task-generation. Execute hard task generation as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-288 EVOLUTION-LAB-SKILL-MUTATION

MH-288 - evolution-lab-skill-mutation. Execute skill mutation as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-290 EVOLUTION-LAB-ROUTER-WEIGHT-OPTIMIZATION

MH-290 - evolution-lab-router-weight-optimization. Execute router weight optimization as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-292 EVOLUTION-LAB-ADVERSARIAL-TASKS

MH-292 - evolution-lab-adversarial-tasks. Execute adversarial tasks as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-287 EVOLUTION-LAB-FAILURE-REPLAY

MH-287 - evolution-lab-failure-replay. Execute failure replay as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-289 EVOLUTION-LAB-PROMPT-VARIANTS

MH-289 - evolution-lab-prompt-variants. Execute prompt variants as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-291 EVOLUTION-LAB-EVALUATION-SCENARIOS

MH-291 - evolution-lab-evaluation-scenarios. Execute evaluation scenarios as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-293 EVOLUTION-LAB-SKILL-ABLATION

MH-293 - evolution-lab-skill-ablation. Execute skill ablation as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

OPERATIONAL EVIDENCE MATRIX

EVOLUTION_LAB

Trigger	Name the condition that activates EVOLUTION_LAB; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Evolution Lab - Skills 1 to 8.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Evolution Lab - Skills 9 to 15

CATEGORY ROLE

Evolution Lab contains auditable capabilities inside the Moussaif Hermes AGI-Oriented Cognitive Kernel. Skills are activated by trigger and contract, not loaded wholesale. Category expansion: category role is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-294 EVOLUTION-LAB-CROSS-VERSION-BENCHMARK

MH-294 - evolution-lab-cross-version-benchmark. Execute cross version benchmark as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-296 EVOLUTION-LAB-SKILL-PROMOTION

MH-296 - evolution-lab-skill-promotion. Execute skill promotion as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-298 EVOLUTION-LAB-EVOLUTION-BUDGET

MH-298 - evolution-lab-evolution-budget. Execute evolution budget as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-300 EVOLUTION-LAB-EVOLUTION-REPORT

MH-300 - evolution-lab-evolution-report. Execute evolution report as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

كل HyperSkill يحتاج trigger وtyped output وfailure test؛ لا يتم تحميل library كاملة لكل task.

ROUTING DISCIPLINE

Select one primary skill, zero to three supports, three to seven knowledge units and one bounded HyperLoop when iteration is necessary. High-risk routes require an independent verifier or court. Category expansion: routing discipline is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

MH-295 EVOLUTION-LAB-REGRESSION-REJECTION

MH-295 - evolution-lab-regression-rejection. Execute regression rejection as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-297 EVOLUTION-LAB-ADAPTER-EVALUATION

MH-297 - evolution-lab-adapter-evaluation. Execute adapter evaluation as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

MH-299 EVOLUTION-LAB-SELF-MODIFICATION-SANDBOX

MH-299 - evolution-lab-self-modification-sandbox. Execute self modification sandbox as an auditable capability within the Evolution Laboratory layer of the Moussaif Hermes AGI-Oriented Cognitive Kernel. Risk=high. Required evidence: valid trigger, schema-valid output, postcondition, telemetry and a negative activation case. The skill card also declares required knowledge, tool permissions, expected artifact, verifier, timeout, stop condition, telemetry fields, memory policy and rollback behavior. Promotion requires one positive, one negative, one edge and one adversarial trajectory; transfer is tested under a different task formulation or model.

CATEGORY PROMOTION GATE

No skill is promoted merely because it is documented. It must pass representative, adversarial and transfer tests, expose failure behavior and remain independently disableable. Category expansion: category promotion gate is connected to the goal contract, sparse router, bounded loop engine and evidence ledger. The category is measured by activation precision, reusable coverage, duplicate rate, transfer, failure containment and the cost of unnecessary loading.

OPERATIONAL EVIDENCE MATRIX

EVOLUTION_LAB

Trigger	Name the condition that activates EVOLUTION_LAB; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Evolution Lab - Skills 9 to 15.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 01 - Theory Admission

CORE RULE

A new theory enters the registry only with a unique problem, causal claim, operational object, dependencies, falsifier, evidence grade and owner. Similarity search against existing theories is mandatory. Admission states are CANDIDATE, PROPOSED, TESTABLE, PARTIALLY SUPPORTED, REJECTED or MERGED. Research note for Protocol Library 01 - Theory Admission: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Protocol Library 01 - Theory Admission: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Protocol Library 01 - Theory Admission: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Protocol Library 01 - Theory Admission: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Protocol Library 01 - Theory Admission: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Protocol Library 01 - Theory Admission: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Protocol Library 01 - Theory Admission: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 01 - Theory Admission.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 02 - Capability Card

CORE RULE

Every capability card defines trigger, inputs, authority, knowledge, algorithm, output schema, verification, telemetry, failure signatures, rollback and memory policy. Capability is demonstrated by positive, negative, edge, adversarial and transfer cases. Research note for Protocol Library 02 - Capability Card: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Protocol Library 02 - Capability Card: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Protocol Library 02 - Capability Card: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Protocol Library 02 - Capability Card: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Protocol Library 02 - Capability Card: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Protocol Library 02 - Capability Card: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Protocol Library 02 - Capability Card: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 02 - Capability Card.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 03 - HyperLoop Contract

CORE RULE

A loop declares an evidence-gain objective, maximum iterations, token and time budgets, saturation threshold, escalation rule and stop predicate. Loop state must be inspectable, replayable and interruptible. Research note for Protocol Library 03 - HyperLoop Contract: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Protocol Library 03 - HyperLoop Contract: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Protocol Library 03 - HyperLoop Contract: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Protocol Library 03 - HyperLoop Contract: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Protocol Library 03 - HyperLoop Contract: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Protocol Library 03 - HyperLoop Contract: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Protocol Library 03 - HyperLoop Contract: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 03 - HyperLoop Contract.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 04 - Court Independence

CORE RULE

Generation and review use separated assumptions, prompts, evidence views and verdict roles. A court may demand remediation, block promotion or require human review. Correlated failures are measured rather than assumed away. Implementation note for Protocol Library 04 - Court Independence: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Protocol Library 04 - Court Independence: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Protocol Library 04 - Court Independence: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Protocol Library 04 - Court Independence: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Protocol Library 04 - Court Independence: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Protocol Library 04 - Court Independence: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Protocol Library 04 - Court Independence: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 04 - Court Independence.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 05 - Memory Write

CORE RULE

A memory write is admitted only if validated, useful, scoped, non-duplicative and privacy-compatible. It carries provenance, owner, expiry, conflict behavior and deletion semantics. Canonical promotion is slower than episodic capture. Governance note for Protocol Library 05 - Memory Write: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Protocol Library 05 - Memory Write: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Protocol Library 05 - Memory Write: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Protocol Library 05 - Memory Write: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Protocol Library 05 - Memory Write: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Protocol Library 05 - Memory Write: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Protocol Library 05 - Memory Write: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 05 - Memory Write.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 06 - Tool Invocation

CORE RULE

Tools expose typed inputs, explicit permissions, preconditions, timeouts, idempotency, postconditions and rollback. Returned state is authoritative; intent or call issuance is never treated as proof of external success. Governance note for Protocol Library 06 - Tool Invocation: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Protocol Library 06 - Tool Invocation: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Protocol Library 06 - Tool Invocation: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Protocol Library 06 - Tool Invocation: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Protocol Library 06 - Tool Invocation: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Protocol Library 06 - Tool Invocation: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Protocol Library 06 - Tool Invocation: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 06 - Tool Invocation.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 07 - Deployment

CORE RULE

Deployment uses adapters, canaries, health probes, SLOs, observability, regression gates, rollback rehearsal and promotion records. Overlays are preferred when they reduce irreversible migration risk. Governance note for Protocol Library 07 - Deployment: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Protocol Library 07 - Deployment: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Protocol Library 07 - Deployment: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Protocol Library 07 - Deployment: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Protocol Library 07 - Deployment: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Protocol Library 07 - Deployment: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Protocol Library 07 - Deployment: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 07 - Deployment.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 08 - Security

CORE RULE

Least privilege, taint tracking, prompt-injection isolation, secret references, data minimization and approval gates apply at tool boundaries. Security evidence includes adversarial traces and containment results. Implementation note for Protocol Library 08 - Security: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Protocol Library 08 - Security: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Protocol Library 08 - Security: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Protocol Library 08 - Security: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Protocol Library 08 - Security: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Protocol Library 08 - Security: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Protocol Library 08 - Security: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 08 - Security.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 09 - Research

CORE RULE

Every study declares baseline, hypothesis, dataset, hidden test, ablation, transfer setting, metric, statistical uncertainty, negative result policy and replication package. Exploratory results are labeled accordingly. Implementation note for Protocol Library 09 - Research: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Protocol Library 09 - Research: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Protocol Library 09 - Research: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Protocol Library 09 - Research: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Protocol Library 09 - Research: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Protocol Library 09 - Research: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Protocol Library 09 - Research: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 09 - Research.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Protocol Library 10 - Audit

CORE RULE

Audit verifies version lineage, artifact integrity, coverage, authority, cost, evidence calibration, data sovereignty, rollback readiness and unresolved risks. A release verdict cannot exceed the weakest critical gate. Implementation note for Protocol Library 10 - Audit: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Protocol Library 10 - Audit: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Protocol Library 10 - Audit: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Protocol Library 10 - Audit: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Protocol Library 10 - Audit: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Protocol Library 10 - Audit: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Protocol Library 10 - Audit: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Protocol Library 10 - Audit.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Benchmark Architecture 01 - Baselines and Ablations

CORE RULE

Compare the full system to ordinary prompting, single-agent execution, uniform routing, no-court verification, no-memory operation and fixed-model baselines. Component value is the causal difference under matched budget and task distribution. Governance note for Benchmark Architecture 01 - Baselines and Ablations: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Benchmark Architecture 01 - Baselines and Ablations: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Benchmark Architecture 01 - Baselines and Ablations: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Benchmark Architecture 01 - Baselines and Ablations: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Benchmark Architecture 01 - Baselines and Ablations: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Benchmark Architecture 01 - Baselines and Ablations: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Benchmark Architecture 01 - Baselines and Ablations: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Benchmark Architecture 01 - Baselines and Ablations.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Benchmark Architecture 02 - Scenario Families

CORE RULE

Use clean, ambiguous, adversarial, constrained, long-horizon and transfer scenarios. Hidden variants prevent prompt memorization. Consequential tasks require postcondition and world-state verification. Implementation note for Benchmark Architecture 02 - Scenario Families: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Benchmark Architecture 02 - Scenario Families: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Benchmark Architecture 02 - Scenario Families: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Benchmark Architecture 02 - Scenario Families: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Benchmark Architecture 02 - Scenario Families: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Benchmark Architecture 02 - Scenario Families: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Benchmark Architecture 02 - Scenario Families: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Benchmark Architecture 02 - Scenario Families.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Benchmark Architecture 03 - Metrics

CORE RULE

Measure correctness, calibration, evidence completeness, cost, latency, token use, route sparsity, tool success, unauthorized action, recovery time, human control, realized value and residual risk. Governance note for Benchmark Architecture 03 - Metrics: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Benchmark Architecture 03 - Metrics: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Benchmark Architecture 03 - Metrics: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Benchmark Architecture 03 - Metrics: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Benchmark Architecture 03 - Metrics: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Benchmark Architecture 03 - Metrics: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Benchmark Architecture 03 - Metrics: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Benchmark Architecture 03 - Metrics.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Benchmark Architecture 04 - Falsification Matrix

CORE RULE

For each theory, construct a positive case, a negative activation case, a counterexample, a removal ablation and an alternative explanation. A theory is revised when its predicted causal difference fails to appear. Governance note for Benchmark Architecture 04 - Falsification Matrix: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Benchmark Architecture 04 - Falsification Matrix: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Benchmark Architecture 04 - Falsification Matrix: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Benchmark Architecture 04 - Falsification Matrix: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Benchmark Architecture 04 - Falsification Matrix: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Benchmark Architecture 04 - Falsification Matrix: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Benchmark Architecture 04 - Falsification Matrix: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Benchmark Architecture 04 - Falsification Matrix.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Benchmark Architecture 05 - Evidence Grades

CORE RULE

E0 is an idea; E1 a specified mechanism; E2 a runnable prototype; E3 internal controlled evidence; E4 independent replication; E5 robust multi-domain support. Marketing, length and fluency do not raise evidence grade. Governance note for Benchmark Architecture 05 - Evidence Grades: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Benchmark Architecture 05 - Evidence Grades: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Benchmark Architecture 05 - Evidence Grades: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Benchmark Architecture 05 - Evidence Grades: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Benchmark Architecture 05 - Evidence Grades: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Benchmark Architecture 05 - Evidence Grades: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Benchmark Architecture 05 - Evidence Grades: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Benchmark Architecture 05 - Evidence Grades.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Implementation Roadmap 01 - Kernel v0.1

CORE RULE

Implement schemas, task contract, world state, sparse router, 30 executable skills, 10 HyperAbilities, bounded-loop engine, HAKIM runner, evidence ledger, courts, verifier, telemetry, replay and benchmark pack. Governance note for Implementation Roadmap 01 - Kernel v0.1: core rule is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Research note for Implementation Roadmap 01 - Kernel v0.1: quality gate declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Governance note for Implementation Roadmap 01 - Kernel v0.1: integration is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Implementation note for Implementation Roadmap 01 - Kernel v0.1: implementation consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Governance note for Implementation Roadmap 01 - Kernel v0.1: required artifact is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Governance note for Implementation Roadmap 01 - Kernel v0.1: failure test is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Research note for Implementation Roadmap 01 - Kernel v0.1: research consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Implementation Roadmap 01 - Kernel v0.1.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Implementation Roadmap 02 - Sovereign Runtime

CORE RULE

Add local model routing, RAG/GraphRAG, SQLite FTS5/BM25 plus vectors, adapters, MCP tool contracts, secrets mapping, Docker/VPS deployment, reverse proxy, TLS, health checks, canaries and rollback. Implementation note for Implementation Roadmap 02 - Sovereign Runtime: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Implementation Roadmap 02 - Sovereign Runtime: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Implementation Roadmap 02 - Sovereign Runtime: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Implementation Roadmap 02 - Sovereign Runtime: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Implementation Roadmap 02 - Sovereign Runtime: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Implementation Roadmap 02 - Sovereign Runtime: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Implementation Roadmap 02 - Sovereign Runtime: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Implementation Roadmap 02 - Sovereign Runtime.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Implementation Roadmap 03 - Research Validation

CORE RULE

Run MCD/Q-MCD ablations, cross-model transfer, context compaction studies, HyperDeployment validation, security red-team, calibration studies and human-agency evaluations before strong AGI claims. Research note for Implementation Roadmap 03 - Research Validation: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Implementation Roadmap 03 - Research Validation: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Implementation Roadmap 03 - Research Validation: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Implementation Roadmap 03 - Research Validation: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Implementation Roadmap 03 - Research Validation: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Implementation Roadmap 03 - Research Validation: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Implementation Roadmap 03 - Research Validation: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Implementation Roadmap 03 - Research Validation.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Publication Agenda

CORE RULE

Separate architecture papers, theory papers, benchmark reports, negative results, software artifacts and integrative books. Each publication names claim status, source lineage, contribution boundary and replication materials. Implementation note for Publication Agenda: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Publication Agenda: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Publication Agenda: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Publication Agenda: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Publication Agenda: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Publication Agenda: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Publication Agenda: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Publication Agenda.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Collaboration Map

CORE RULE

Collaboration can target consciousness theory, computational cybernetics, agentic systems, formal methods, HCI, software reliability, causal inference, benchmarking or governance. The best first ask is a bounded co-evaluation rather than endorsement of the whole corpus. Research note for Collaboration Map: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Collaboration Map: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Collaboration Map: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Collaboration Map: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Collaboration Map: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Collaboration Map: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Collaboration Map: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Collaboration Map.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Canonical Crosswalk

CORE RULE

MUT unifies theory; MUG governs the program; M-AGIOS-K is the kernel theory; MOUSSAIF AGI OS is the architecture; MOUSSAIXT is the execution shell; MHKOS is knowledge-as-OS; Q-MCD governs questions; MCD ranks causal differences; MOFIP is the open Fattah research program. Implementation note for Canonical Crosswalk: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for Canonical Crosswalk: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for Canonical Crosswalk: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for Canonical Crosswalk: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for Canonical Crosswalk: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for Canonical Crosswalk: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for Canonical Crosswalk: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Canonical Crosswalk.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Glossary and Status Language

CORE RULE

Use theory, hypothesis, mechanism, architecture, prototype, runtime, tool, skill, benchmark, corpus and product precisely. Proposed means internally formulated; supported means evidence exists; validated requires the declared threshold; deployed means operational in a stated environment. Research note for Glossary and Status Language: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Glossary and Status Language: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Glossary and Status Language: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Glossary and Status Language: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Glossary and Status Language: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Glossary and Status Language: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Glossary and Status Language: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Glossary and Status Language.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

External Reference Anchors

CORE RULE

Reference anchors include Turing on machine intelligence; Newell and Simon on physical symbol systems; Baars and Dehaene on global workspace; Tononi on integrated information; Friston on active inference; Pearl on causality; Sutton and Barto on reinforcement learning; Russell and Norvig on AI foundations. Implementation note for External Reference Anchors: core rule maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Governance note for External Reference Anchors: quality gate is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Implementation note for External Reference Anchors: integration maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Research note for External Reference Anchors: implementation consequence declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Implementation note for External Reference Anchors: required artifact maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Implementation note for External Reference Anchors: failure test maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Governance note for External Reference Anchors: research consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by External Reference Anchors.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet

Closing Research Constitution

CORE RULE

Preserve user authority, truthfulness, privacy, safety and explicit constraints. Compile goals; govern questions; distinguish state types; route sparsely; plan reversibly; act through typed tools; observe returned state; verify independently; commit or rollback; learn only validated memory; stop on success, saturation, risk or revoked authority. Research note for Closing Research Constitution: core rule declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

QUALITY GATE

A reviewer must be able to reconstruct the decision, reproduce the test, identify authority and determine whether the artifact changed the route or outcome. Implementation note for Closing Research Constitution: quality gate maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

INTEGRATION

Map the result to MUT for theory integrity, MUG for program governance, M-AGIOS-K for kernel routing and MOUSSAIXT for execution and telemetry. Research note for Closing Research Constitution: integration declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

IMPLEMENTATION CONSEQUENCE

Every consequential mutation uses explicit permission, pre-state capture, postcondition verification and rollback or residual-state reporting. Governance note for Closing Research Constitution: implementation consequence is evaluated at design, execution and release time. Authority, privacy, evidence, reversibility and human control remain visible. The runtime records actual returned state and court verdict. Promotion is denied when the artifact cannot be reproduced, a critical postcondition fails, rollback is untested or the claim exceeds its evidence grade.

الخلاصة: mechanism + artifact + falsifier + evidence grade؛ غير ذلك يبقى concept غير operational.

REQUIRED ARTIFACT

Emit a versioned, schema-valid record with owner, source lineage, evidence pointers, status, residual uncertainty and next action. Research note for Closing Research Constitution: required artifact declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

FAILURE TEST

Construct a case where the rule should refuse activation, a case with conflicting evidence, an adversarial case and a recovery case. Record false positives as well as missed defects. Research note for Closing Research Constitution: failure test declares a falsifiable prediction and the artifact needed to test it. Baseline, ablation, transfer and replication are planned before strong language is used. The outcome is written to the evidence ledger with configuration, limitations and negative findings, then connected to the next bounded experiment rather than converted directly into a portfolio-wide claim.

RESEARCH CONSEQUENCE

Promotion depends on measured causal value under matched baselines. If the mechanism does not improve outcome or safety, simplify, merge or reject it. Implementation note for Closing Research Constitution: research consequence maps to an accountable record, explicit decision boundary and measurable result. The record states assumptions, authority, evidence, version, risk, cost, stop condition and residual uncertainty. Review includes a negative case, adversarial case and recovery case. If marginal value is absent under matched conditions, simplify, merge or reject; if evidence is incomplete, status remains HOLD.

OPERATIONAL EVIDENCE MATRIX

SYNTHESIS

Trigger	Name the condition that activates SYNTHESIS; reject unnecessary activation.	Negative case must route to the simpler baseline.	trigger + reason
Inputs	Goal, state, evidence, authority, budget, schema and version are explicit.	Missing or stale critical input produces HOLD.	input manifest
Object	Emit the typed artifact required by Closing Research Constitution.	Invalid schema or inert artifact fails the mechanism claim.	artifact + hash
Evidence	Attach provenance, grade, uncertainty and alternative explanations.	Counterevidence and source conflict remain visible.	claim ledger
Authority	Check scope, privacy, reversibility and human approval at action time.	Revoked or ambiguous authority blocks mutation.	authority log
Baseline	Match model, tools, data, time and token budget; ablate the component.	No marginal gain leads to simplify, merge or reject.	ablation record
World effect	Declare expected delta, observe returned state and verify postcondition.	Intent or call issuance never proves external success.	state delta
Recovery	Define stop, rollback, residual state and validated memory write.	Failed reversal escalates with preserved evidence.	recovery packet