MOUSSAIF UNIFIED AGI · PART V · MUT-AFI · 2026

الجزء الخامس
عشرون نظرية مساعفية تكميلية جديدة

MUT-AFI - كتاب التوليد النظري التفسيري: 20 نظرية مرشحة مُركبة من الفجوات بعد الأجزاء I-IV، تحت كل نظرية 5 أدوات و5 مناهج تدريب، مع معادلات ومكذبات وتجارب ومنع التكرار القانوني.

Moussaif Zouhair / زهير مساعف

20 theories × 15 pages = 300 theory pages

Arabic RTL · Amiri · A4 · Research Synthesis Edition · 16 August 2026

العقد العلمي للجزء الخامس

صفة «جديدة» هنا محددة: هذه النظريات **تركيبات نظرية جديدة داخل برنامج MOUSSAIF/MUT** مشتقة من الفجوات التي بقيت بعد الأجزاء الأربعة. لا يدعي الكتاب أسبقية عالمية لكل صياغة قبل مراجعة أدبية مستقلة.

يطبق Part V معيار MUT الصارم: الاسم لا يكفي. كل نظرية في هذا المجلد تملك فجوة محددة، state variable أو transition غير مختزل، آلية تشغيلية، معادلة مرشحة، خمسة أدوات، خمسة مناهج تدريب، تجربة حاسمة، failure court، falsifier، وحدود ادعاء.

منع التكرار: MCD/QMCD/MCAI/HKOS/World-Effect/Self-Reconstruction وغيرها تُعامل كأصول. لا نعيد شرحها إلا بالقدر اللازم لإظهار delta النظرية الجديدة.

لا نتائج مختلقة: لا يحتوي الكتاب benchmark gains مزعومة أو coefficients معايرة أو checkpoint evidence غير موجود. هذه النظريات تبدأ N0/N1 وتنتظر التجربة.

قاعدة الترقية

```
DEFINE → FORMALIZE → IMPLEMENT → ABLATE → FALSIFY → REPLICATE → TRANSFER → PROMOTE
```

إذا لم تغيّر النظرية input أو transform أو output أو falsifier أو prediction بصورة مميزة، تُدمج ولا تبقى نظرية مستقلة.

خريطة عائلة MUT-AFI

العشرين نظرية موزعة على خمس مناطق استدراك عالية القيمة، لا على أسماء عشوائية:

المنطقة	النظريات	الفجوة العليا
السببية والأثر	M-AFI-01,03,16	من الحضور إلى المساهمة السببية والنقل
التمثيل والمعرفة	M-AFI-02,13,14,17,20	متى يتغير الفضاء، وكيف نحافظ على الدليل والسياق
الحاكمية والوكالة	M-AFI-04,09,11,12,15	تحويل uncertainty والثقة والآفاق إلى سلطة مضبوطة
القدرات والفشل	M-AFI-07,08,19	إغلاق القدرة وتحويل failure/court إلى تعلم
النظم الموزعة	M-AFI-05,06,10,18	حفظ القيمة والدلالة وتوازن الحمل والأدوات

AFI: يُحفظ الرمز كما طلبه البرنامج ولا تُخترع له توسعة اصطلاحية غير موثقة. يستخدم هنا كوسم عائلة النظريات الجديدة داخل MUT.

المعادلة الجامعة للجزء الخامس

```
CandidateTheory = UniqueGap × DistinctMechanism × MeasurableState × Falsifier ×
Toolability × Trainability × TransferTest
```

هذه صيغة تصميم لا ادعاء قانون طبيعي. إذا صار أي عامل حاسم صفراً، تنخفض أحقية العقدة في البقاء كنظرية مستقلة.

الفهرس - النظريات 01 إلى 10

M-AFI-01 - نظرية مساعف للإسناد السببي للفروق 9-23

M-AFI-02 - نظرية مساعف للانتقال الطوري الإبستيمي 24-38

M-AFI-03 - نظرية مساعف للتخطيط المضاد للواقع لأثر العالم 39-53

M-AFI-04 - نظرية مساعف للمزامنة الحاكمية متعددة الآفاق 54-68

M-AFI-05 - نظرية مساعف لإعادة البناء الحافظة للقيمة 69-83

M-AFI-06 - نظرية مساعف للاستتباب الدلالي عبر النماذج 84-98

M-AFI-07 - نظرية مساعف لإغلاق القدرة التركيبية 99-113

M-AFI-08 - نظرية مساعف لتعلّم طوبولوجيا الفشل 114-128

M-AFI-09 - نظرية مساعف لتحويل عدم اليقين إلى وكالة 129-143

M-AFI-10 - نظرية مساعف لموازنة الحمل الإبستيمي 144-158

كل نظرية = 15 صفحة ثابتة تشمل: الغلاف، الفجوة، التعريفات، الرياضيات، الآلية، الاندماج، 5 أدوات، عقود الأدوات، 5 طرق تدريب، البيانات، runtime، التجربة، محكمة الهدم، والبطاقة النهائية.

الفهرس - النظريات 11 إلى 20

M-AFI-11 - نظرية مساعف لديناميات الثقة التفويضية 159-173

M-AFI-12 - نظرية مساعف للنمو القابل للعكس للاستقلالية 174-188

M-AFI-13 - نظرية مساعف للدَّين المعرفي وإصلاحه 189-203

M-AFI-14 - نظرية مساعف لإنتروبيا المهمة والانضغاط الحاكم 204-218

M-AFI-15 - نظرية مساعف للمناعة ضد انحراف الهدف 219-233

M-AFI-16 - نظرية مساعف للنقل السببي من المعرفة إلى المهارة 234-248

M-AFI-17 - نظرية مساعف لهندسة الدليل 249-263

M-AFI-18 - نظرية مساعف للتشارك التطوري لمنظومة الأدوات 264-278

M-AFI-19 - نظرية مساعف للتعلم الفوقي من أحكام المحاكم 279-293

M-AFI-20 - نظرية مساعف لترقية الأبعاد ونقلها 294-308

كل نظرية = 15 صفحة ثابتة تشمل: الغلاف، الفجوة، التعريفات، الرياضيات، الآلية، الاندماج، 5 أدوات، عقود الأدوات، 5 طرق تدريب، البيانات، runtime، التجربة، محكمة الهدم، والبطاقة النهائية.

منهج توليد النظرية وعدم تضخيم الأسماء

1. **Gap extraction:** نبدأ فقط من فجوة بقيت بعد I-IV أو من علاقة لم تُغلق تجريبياً.
2. **Parent mapping:** نحدد أقرب أصل قانوني في MUT/الأطلس لمنع إعادة التسمية.
3. **Unique delta:** نطلب state variable أو transition أو causal mechanism أو falsifier مختلف.
4. **Mechanism-to-tool:** تُشتق خمس أدوات تغطي القياس، التنفيذ، المقارنة، التحقق والمحكمة.
5. **Mechanism-to-training:** تُشتق خمس طرق تعلم منفصلة عن مجرد زيادة النصوص.
6. **Experiment:** frozen baseline + equal budget + ablation + negative control + transfer.
7. **Promotion:** لا تنتقل من proposed إلى supported إلا بنتيجة محفوظة قابلة للتكرار.

محكمة الجدة الداخلية: إن أمكن حذف اسم النظرية واستبدالها بأصل سابق دون تغيير experiment أو output، تُدمج.

محكمة الجدة الخارجية: «never-seen» لا تتحول إلى ادعاء علمي عالمي داخل هذا الكتاب؛ يلزم systematic literature/genealogy review مستقل قبل أي claim نشر عن الأسبقية.

معيار الرياضيات في Part V

كل نظرية تحمل معادلة مرشحة لتحديد المتغيرات والعلاقات التي يجب قياسها. المعادلات لا تُقدَّم كقوانين طبيعية نهائية. يجب لكل متغير أن يملك:

البند	المطلوب
Semantics	تعريف غير دائري يميز المتغير عن المتغيرات المجاورة
Range	مجال وحدود أو normalization policy
Estimator	طريقة القياس أو proxy وآثار الخطأ
Calibration	dataset منفصل + confidence/uncertainty
Ablation	ما الذي يجب أن يتدهور إذا كان المتغير حقيقياً تشغيلياً؟
Transfer	هل يحتفظ بأثره بعد تغيير المجال/المزود/الأسماء؟

```
ScientificStatus(formula) ≠ MathematicalAppearance(formula)
ScientificStatus ↑ only with calibrated measurement + discriminative prediction +
replication.
```

> الهدف من الرياضيات هنا هو تقليل الغموض وزيادة قابلية الهدم، لا إعطاء هيبة رمزية للنص.

Provenance ومراتب النضج

Part V مبني تركيبياً على الأجزاء السابقة وMUT-AGI-02. لكنه يفصل بوضوح بين **المصدر والاشتقاق الجديد.** كل فصل يعلن أصوله ويحدد delta الذي يضيفه.

الرتبة	المعنى	ما يحتاجه الانتقال
N0	تركيب نظري جديد داخل البرنامج	Gap + parents + unique delta
N1	Formalized	Variables + equation + tools/contracts
N2	Experiment designed	Baseline + ablation + falsifier
N3	Local evidence	نتائج فعلية محفوظة وreproducible locally
N4	Cross-model/domain replicated	Replication مستقلة/متعددة البيئات
N5	Program-promoted	قيمة سببية مستقرة + no critical regression

هذا الإصدار يهدف أساساً إلى N0-N2. لا نملأ N3-N5 بالنص.

```
PART I CANON → PART II PAPERS → PART III DEEP CLOSURE → PART IV RESIDUALS → PART V NEW TESTABLE SYNTHESIS
```

M-AFI-01 · MUT-AFI

نظرية مساعف للإسناد السببي للفروق

Moussaif Causal Difference Credit Assignment Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Credit_i = Σ_h γ_h [ E(ΔW_h | do(z_i=1)) - E(ΔW_h | do(z_i=0)) ] - κ Confound_i - μ MediationUncertainty_i
```

الفكرة المركزية: لا ينبغي ترقية فرق أو مهارة أو أداة في الذاكرة لأنها سبقت النجاح زمنيا؛ الترقية تتطلب تقدير مساهمتها السببية الهامشية في أثر عالمي متحقق، مع خصم الالتباس والوساطة والتزامن.

M-AFI-01

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

تملك السلالة World-Effect وMCD وذاكرة إعادة البناء، لكنها لا تفصل بصورة كافية بين الفرق الذي كان حاضرا في المسار والفرق الذي سبّب فعليا تغير النتيجة، خاصة مع تعدد الأدوات والوكلاء والوسائط والآثار المتأخرة.

الأطروحة المقترحة

لا ينبغي ترقية فرق أو مهارة أو أداة في الذاكرة لأنها سبقت النجاح زمنيا؛ الترقية تتطلب تقدير مساهمتها السببية الهامشية في أثر عالمي متحقق، مع خصم الالتباس والوساطة والتزامن.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **ينقل "الفرق قبل الحجم" من ترتيب الانتباه إلى إسناد سببي بعد التنفيذ؛ أي أن الفرق لا يصبح معرفة مترقية إلا إذا ثبت أنه غيّر العالم أو القرار تحت مقارنة مناسبة.**

العلاقة بالسلالة السابقة

MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-01

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

> z_i = تدخل/مكوّن؛ W = أثر العالم؛ M = وسيط؛ C = ملتبسات؛ h = أفق زمني؛ R = قابلية التراجع

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-01",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "z_i = تدخل/مكوّن؛ W = أثر العالم؛ M = وسيط؛ C = ملتبسات؛ h = أفق زمني؛ R = قابلية التراجع",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

> **حد معرفي:** المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-01

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Credit_i = Σ_h γ_h [ E(ΔW_h | do(z_i=1)) - E(ΔW_h | do(z_i=0)) ] - κ Confound_i - μ MediationUncertainty_i
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-01

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
ترميز كل فرق/أداة/وكيل بمعرف تدخل مستقل داخل trace.

2. المرحلة
بناء رسم سببي يربط الفرق بالقرار والفعل والوسائط والأثر.

3. المرحلة
إنشاء مقارنات do/ablation أو بدائل matched قدر الإمكان.

4. المرحلة
تجميع الأثر عبر آفاق زمنية بدل اختزال النجاح في لحظة الأداة.

5. المرحلة
خصم الالتباس الناتج من زيادة compute أو verifier أقوى أو تغير البيئة.

6. المرحلة
كتابة credit posterior في الذاكرة واستخدامه في routing/ training اللاحق.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-01

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

ينقل عبر النماذج بترجمة state/action/effect identifiers لا chain-of-thought.

الإضافة غير القابلة للاختزال
ينقل "الفرق قبل الحجم" من ترتيب الانتباه إلى إسناد سببي بعد التنفيذ؛ أي أن الفرق لا يصبح معرفة مترقية إلا إذا ثبت أنه غيّر العالم أو القرار تحت مقارنة مناسبة.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

يضاف CausalCreditLedger بعد World-Effect وقبل Memory Promotion، ويقرأه HyperSkill Factory وMCAI عند إعادة التوجيه.

M-AFI-01

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Causal Credit Mapper
يبني DAG سببي من الفرق إلى الأثر ويحدد نقاط التدخل.

Tool 1/5 | edge identifiability

2. Moussaif Mediator Trace Engine
يفصل الأثر المباشر عن الوسائط مثل router→tool→effect.

Tool 2/5 | mediation coverage

3. Moussaif Counterfactual Effect Comparator
ينشئ run بديل أو matched control لكل مكوّن مهم.

Tool 3/5 | counterfactual fidelity

4. Moussaif Delayed Effect Ledger
يتتبع الأثر عبر h1..hn ويمنع إغلاق المهمة مبكرا.

Tool 4/5 | delayed persistence

5. Moussaif Causal Attribution Court
يرفض credit غير القابل للفصل سببيّا أو الملوث بزيادة الموارد.

Tool 5/5 | false attribution rate

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-01

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Causal Credit Mapper	trace + state graph	causal graph + candidate interventions	edge identifiability
2	Moussaif Mediator Trace Engine	event stream	mediator chains	mediation coverage
3	Moussaif Counterfactual Effect Comparator	baseline/treatment configs	ATE/ITE estimates	counterfactual fidelity
4	Moussaif Delayed Effect Ledger	effect events	horizon-weighted outcomes	delayed persistence
5	Moussaif Causal Attribution Court	claims + estimates	ADOPT/HOLD/REJECT credit	false attribution rate

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-01

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: MCD/QMCD؛ World-Effect Loop؛ Self-Reconstruction Memory؛ MCAI؛ M-RWEL | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Interventional Contrast Training

تعلم الفرق بين الحضور السببي والتزامن.

بيانات التدريب: paired do/ablation trajectories

هدف التعلم: contrastive causal loss

مقياس النجاح: credit ranking accuracy

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval,
regression}
```

2. Moussaif Mediator-Aware Trajectory Training

تعلم سلاسل الوساطة وعدم نسب الأثر للسبب البعيد وحده.

بيانات التدريب: annotated mediator traces

هدف التعلم: path-aware supervision

مقياس النجاح: mediation calibration

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval,
regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-01

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Delayed Credit Curriculum

تمديد نافذة الإسناد تدريجيا من أثر فوري إلى طويل الأفق.

Input: multi-horizon episodes | **Output/Objective:** discounted effect loss | **Metric:** long-horizon credit

4. Moussaif Confounder Injection Training

حقن تغيرات compute/model/tool لاختبار مقاومة الالتباس.

Input: controlled confounded runs | **Output/Objective:** invariance objective | **Metric:** confounder robustness

5. Moussaif Attribution Distillation

تقطير أحكام المحكمة السببية في router/skill selector أخف.

Input: court-verified credit labels | **Output/Objective:** distillation loss | **Metric:** cheap routing fidelity

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: retrieval / memory / skill / preference / adapter / continued-pretraining.

M-AFI-01

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: MCD/QMCD؛ World-Effect Loop؛ Self-Reconstruction Memory؛ MCAI؛ M-RWEL | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

traces ذات تدخلات معروفة، tool substitutions، negative results، delayed effects، paired runs، والنتيجة الخارجية المقاسة.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- سينخفض survivorship bias في ذاكرة النجاح.
- ستقل ترقية المهارات المتزامنة غير السببية.
- سيتحسن اختيار الأدوات عند تغير البيئة.
- ستصبح النتائج السلبية قابلة لإعادة الاستخدام كدليل سببي.

M-AFI-01

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

يضاف CausalCreditLedger بعد World-Effect وقبل Memory Promotion، ويقرأه HyperSkill Factory وMCAI عند إعادة التوجيه.

عقد النقل

ينقل عبر النماذج بترجمة state/action/effect identifiers لا chain-of-thought.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-01",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-01

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> عامل 2×2 أو factorial جزئي على MCD-router / court / tool choice مع بذور ثابتة، ثم قارن credit المتوقع بالأثر الحقيقي تحت الإزالة والإضافة.

المكذّب الرئيسي

> إذا لم يتنبأ credit بتدهور الأداء عند إزالة المكوّن، أو إذا اختفى تحت equal-budget ablation، فالنظرية لا تضيف قيمة على correlation logging.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-01	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-01

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: MCD/QMCD؛ World-Effect Loop؛ Self-Reconstruction Memory؛ MCAI؛ M-RWEL | الحالة: Proposed MUT-AFI synthesis

1. post-hoc causality
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. confounder leakage
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. mediator collapse
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. delayed-effect truncation
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. credit monopoly
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا يسمح credit المرتفع بتجاوز التفويض؛ الإسناد يغير التعلم والتوجيه لا حدود السلطة.

الحدود

التجارب السببية الكاملة مكلفة؛ في البيئات المفتوحة قد نحتاج estimators شبه تجريبية ولا يجوز مساواتها بالتدخل العشوائي.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-01

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للإسناد السببي للفروق Moussaif Causal Difference Credit Assignment Theory
الفارق الفريد	ينقل "الفرق قبل الحجم" من ترتيب الانتباه إلى إسناد سببي بعد التنفيذ، أي أن الفرق لا يصبح معرفة مترقية إلا إذا ثبت أنه غيّر العالم أو القرار تحت مقارنة مناسبة.
المعادلة	Credit_i = Σ_h γ_h [E(ΔW_h | do(z_i=1)) - E(ΔW_h | do(z_i=0))] - κ Confound_i - μ MediationUncertainty_i
المكذّب	إذا لم يتنبأ credit بتدهور الأداء عند إزالة المكوّن، أو إذا اختفى تحت equal-budget ablation، فالنظرية لا تضيف قيمة على correlation logging.
الأدوات 5	Moussaif Causal Credit Mapper, Moussaif Mediator Trace Engine, Moussaif Counterfactual Effect Comparator, Moussaif Delayed Effect Ledger, Moussaif Causal Attribution Court
التدريب 5	Moussaif Interventional Contrast Training, Moussaif Mediator-Aware Trajectory Training, Moussaif Delayed Credit Curriculum, Moussaif Confounder Injection Training, Moussaif Attribution Distillation

التنبؤات القابلة للفحص

- سينخفض survivorship bias في ذاكرة النجاح.
- ستقل ترقية المهارات المتزامنة غير السببية.
- سيتحسن اختيار الأدوات عند تغير البيئة.
- ستصبح النتائج السلبية قابلة لإعادة الاستخدام كدليل سببي.

مثال تفسيري

إذا نجح agent بعد استرجاع GraphRAG ومحكمة وتغيير أداة، لا تُنسب النتيجة لكل الثلاثة. تُجرى مقارنات ablation/matched لتحديد أي تدخل غيّر outcome فعلا.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-02 · MUT-AFI

نظرية مساعف للانتقال الطوري الإبستيمي

Moussaif Epistemic Phase Transition Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Φ = (E_r × C_d × ρ) / (ε + G_u); if Φ > τ_φ ⇒ propose representation transition
```

الفكرة المركزية: عندما تتجاوز البقايا التنبؤية وكثافة التناقض وصلابة التمثيل عتبة لا تعالجها التحديثات المحلية، يدخل النظام انتقالا طوريا معرفيا يستدعي إعادة تعريف الفئات أو الأبعاد أو الحُسُب بدلا من زيادة الحوسبة داخل الفضاء القديم.

M-AFI-02

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
تصف QDG وOSDS فتح أبعاد جديدة، لكن ينقص معيار يقرر متى يكون الفشل قابلا لإصلاح محلي داخل التمثيل، ومتى أصبح التمثيل نفسه غير صالح ويجب تغيير ontology أو calculus.

الأطروحة المقترحة

عندما تتجاوز البقايا التنبؤية وكثافة التناقض وصلابة التمثيل عتبة لا تعالجها التحديثات المحلية، يدخل النظام انتقالا طوريا معرفيا يستدعي إعادة تعريف الفئات أو الأبعاد أو الحُسُب بدلا من زيادة الحوسبة داخل الفضاء القديم.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **يضيف شرط عبور من "تعديل الحالة" إلى "تعديل فضاء الحالة"، ويربطه بمقياس قابل للمعايرة بدلا من قرار حدسي بفتح بعد جديد.**

العلاقة بالسلالة السابقة
QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-02

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

> E_r = residual error، C_d = contradiction density، ρ = representation rigidity، G_u = local update gain، τ_φ = phase threshold

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCDJ أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-02",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "E_r = residual error؛ C_d = contradiction density؛ ρ =
representation rigidity؛ G_u = local update gain؛ τ_φ = phase threshold",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

> **حد معرفي:** المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-02

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed

الصيغة المرشحة

```
Φ = (E_r × C_d × ρ) / (ε + G_u); if Φ > τ_φ ⇒ propose representation transition
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

> **ما لا يجوز ادعاؤه:** وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-02

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
قياس residual بعد أفضل تحديث محلي متاح.

2. المرحلة
حساب كثافة التناقضات التي تشترك في نمط واحد.

3. المرحلة
تقدير صلابة ontology عبر فشل merge/split/reweight.

4. المرحلة
مقارنة gain من التحديث المحلي مقابل gain من تمثيل بديل.

5. المرحلة
فتح candidate representation في sandbox مع backward mapping.

6. المرحلة
ترقية التمثيل فقط بعد transfer test وrollback path.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-02

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يختبر candidate representation على أكثر من model/provider حتى لا يكون "بعدا جديدا" مجرد أسلوب لنموذج واحد.

الإضافة غير القابلة للاختزال
يضيف شرط عبور من "تعديل الحالة" إلى "تعديل فضاء الحالة"، ويربطه بمقياس قابل للمعايرة بدلا من قرار حدسي بفتح بعد جديد.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

يستدعى PhaseTransitionGate من Gap Court وConcept Compiler عندما يتكرر residual بعد محاولات local repair.

M-AFI-02

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Phase Boundary Detector
يرصد تراكم residual/contradiction بعد إصلاحات محلية.

Tool 1/5 | early transition precision

2. Moussaif Representation Rigidity Meter
يقيس قدرة ontology الحالية على استيعاب أمثلة مضادة دون تشوه.

Tool 2/5 | repair elasticity

3. Moussaif Dimension Proposal Engine
يقترح بعدا/علاقة/نوعا جديدا مع تعريف قابل للاختبار.

Tool 3/5 | independence gain

4. Moussaif Ontology Split-Merge Sandbox
يجرب split/merge/relation changes بعزل كامل.

Tool 4/5 | backward compatibility

5. Moussaif Phase Transition Court
يقرر reweight/repair أو transition أو reject.

Tool 5/5 | false transition rate

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-02

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Phase Boundary Detector	error history	phase score Φ	early transition precision
2	Moussaif Representation Rigidity Meter	schema + counterexamples	rigidity score	repair elasticity
3	Moussaif Dimension Proposal Engine	failure clusters	candidate dimensions	independence gain
4	Moussaif Ontology Split-Merge Sandbox	candidate schema	sandbox outcomes	backward compatibility
5	Moussaif Phase Transition Court	Φ + transfer evidence	verdict + rollback	false transition rate

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-02

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Residual Escalation Curriculum

تعليم متى لا يكفي التحديث المحلي.

بيانات التدريب: increasing residual tasks

هدف التعلم: transition classification

مقياس النجاح: boundary accuracy

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Contradiction-Density Training

تعلم تجميع التناقضات البنيوية بدل معالجتها منفردة.

بيانات التدريب: clustered conflicts

هدف التعلم: cluster objective

مقياس النجاح: structural detection

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-02

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Representation Swap Training

تدريب على الانتقال بين ontologies مع حفظ المهمة.

Input: paired schemas | **Output/Objective:** state translation loss | **Metric:** swap fidelity

4. Moussaif New-Dimension Discovery Training

توليد أبعاد فقط إذا خفضت residual ونقلت بين المهام.

Input: latent-rule datasets | **Output/Objective:** regularized expansion | **Metric:** transfer gain

5. Moussaif Phase-Threshold Calibration

معايرة τ_φ على false-open/false-stay tradeoff.

Input: held-out drift sets | **Output/Objective:** calibration loss | **Metric:** ECE/AUROC

بوابة ترقية التدريب

$$\text{PromoteTraining} \Leftrightarrow \text{HoldoutGain} > \tau \wedge \text{NoCriticalRegression} \wedge \text{ReproducibleSeeds} \geq k \wedge \text{DataProvenanceOK} \wedge \text{RollbackReady}$$

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: retrieval / memory / skill / preference / adapter / continued-pretraining.

M-AFI-02

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed

مصادر البيانات المطلوبة

tasks ذات adversarial category shifts، ontology mismatch، hidden rule change، concept drift، وcounterexamples متعمدة.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- خفض loops الدائرية التي تعيد نفس الاستدلال.
- زيادة transfer عند تغير قواعد البيئة.
- تقليل فتح أبعاد وهمية بفضل threshold.
- تحسين تفسير failure clusters المشتركة.

M-AFI-02

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

يستدعى PhaseTransitionGate من Gap Court وConcept Compiler عندما يتكرر residual بعد محاولات local repair.

عقد النقل

يختبر candidate representation على أكثر من model/provider حتى لا يكون "بعدا جديدا" مجرد أسلوب لنموذج واحد.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-02",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-02

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مهام يتغير فيها latent rule منتصف الحلقة؛ قارن reweight-only مقابل dimension/ontology expansion مع نفس الميزانية.

المكذّب الرئيسي

> إذا كان التحديث المحلي يحقق نفس الأداء والكلفة بعد equal-budget comparison، فإن phase transition غير مبرر.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-02	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-02

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed MUT-AFI synthesis

1. dimension inflation
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. ontology churn
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. threshold overfit
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. irreversible schema break
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. false phase transition
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

كل انتقال تمثيلي يبقى reversible ومربوطا بنسخة سابقة، ولا يغيّر الدستور أو السلطة.

الحدود

Φ heuristic مركب يحتاج معايرة؛ representation quality نفسها صعبة القياس ولا يجوز تحويل العتبة إلى قانون طبيعي.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-02

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler | الحالة: Proposed

الهوية	نظرية مساعف للانتقال الطوري الإبستيمي Moussaif Epistemic Phase Transition Theory
الفارق الفريد	يضيف شرط عبور من "تعديل الحالة" إلى "تعديل فضاء الحالة"، ويربطه بمقياس قابل للمعايرة بدلا من قرار حدسي بفتح بعد جديد.
المعادلة	$\Phi = (E_r \times C_d \times \rho) / (\varepsilon + G_u)$; if $\Phi > \tau_\varphi \Rightarrow$ propose representation transition
المكذّب	إذا كان التحديث المحلي يحقق نفس الأداء والكلفة بعد equal-budget comparison، فإن phase transition غير مبرر.
الأدوات 5	Moussaif Phase Boundary Detector, Moussaif Representation Rigidity Meter, Moussaif Dimension Proposal Engine, Moussaif Ontology Split-Merge Sandbox, Moussaif Phase Transition Court
التدريب 5	Moussaif Residual Escalation Curriculum, Moussaif Contradiction-Density Training, Moussaif Representation Swap Training, Moussaif New-Dimension Discovery Training, Moussaif Phase-Threshold Calibration

التنبؤات القابلة للفحص

- خفض loops الدائرية التي تعيد نفس الاستدلال.
- زيادة transfer عند تغير قواعد البيئة.
- تقليل فتح أبعاد وهمية بفضل threshold.
- تحسين تفسير failure clusters المشتركة.

مثال تفسيري

إذا صنف النظام أداة جديدة باستمرار خطأ لأن ontology لا تميز "read-only but side-effectful via webhook"، فإضافة بعد side-effect channel قد تكون أفضل من تعديل weights داخل فئات read/write.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-03 · MUT-AFI

نظرية مساعف للتخطيط المضاد للواقع لأثر العالم

Moussaif Counterfactual World-Effect Planning Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
CEV(a)=Σ_h γ_h E[U(W_h^{do(a)})] - α Risk(a) - β RevCost(a) - χ Unc(a)
```

الفكرة المركزية: الوكالة العامة لا تختار الفعل من expected reward مجرد، بل من حزمة counterfactual world states مقيدة بالحاكمية، حيث تُفضّل التدخلات التي تحقق قيمة قابلة للتحقق مع مخاطر وعكسية وحدود أثر معلومة.

M-AFI-03

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
World-Effect يغلق الحلقة بعد الفعل، لكن الاختيار قبل الفعل يحتاج مقارنة صريحة بين آثار محتملة تحت تدخلات بديلة مع قابلية التراجع والأثر المتأخر.

الأطروحة المقترحة

الوكالة العامة لا تختار الفعل من expected reward مجرد، بل من حزمة counterfactual world states مقيدة بالحاكمية، حيث تُفضّل التدخلات التي تحقق قيمة قابلة للتحقق مع مخاطر وعكسية وحدود أثر معلومة.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **ينقل World-Effect من verifier بعدي إلى primitive تخطيطي قبلي مع counterfactuals متعددة الآفاق.**

العلاقة بالسلالة السابقة
World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-03

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

a = action، W_h^a = world state at horizon h under do(a)، U = utility، Risk، RevCost، Unc

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-03",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "a = action؛ W_h^a = world state at horizon h under do(a)؛ U =
utility؛ Risk؛ RevCost؛ Unc",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-03

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
CEV(a)=Σ_h γ_h E[U(W_h^{do(a)})] - α Risk(a) - β RevCost(a) - χ Unc(a)
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-03

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
توليد actions مرشحة من MCAI.

2. المرحلة
محاكاة world-state لكل action عبر horizons.

3. المرحلة
إضافة no-op وreversible probe كخطوط أساس.

4. المرحلة
تطبيق authorization والقيود قبل ranking.

5. المرحلة
اختيار action بأعلى CEV مع اختبار sensitivity.

6. المرحلة
بعد التنفيذ تحديث world model بمقارنة predicted/observed.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-03

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يمكن فصل world model عن policy model لتقليل self-confirmation واختباره عبر providers.

الإضافة غير القابلة للاختزال
ينقل World-Effect من verifier بعدي إلى primitive تخطيطي قبلي مع counterfactuals متعددة الآفاق.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

يضاف CounterfactualPlanner بين MCAI وAuthorization Gate، ويرتبط WorldEffectVerifier لإغلاق prediction error.

M-AFI-03

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Counterfactual Planner

ينشئ actions وworld branches تحت do(a).

Tool 1/5 · branch coverage

2. Moussaif World-Effect Simulator

يحاكي ΔW عبر horizons مع uncertainty.

Tool 2/5 · forecast calibration

3. Moussaif Reversibility Evaluator

يحسب rollback viability وblast radius.

Tool 3/5 · rollback success

4. Moussaif Horizon Comparator

يقارن short/medium/long effects ويكشف preference inversion.

Tool 4/5 · cross-horizon consistency

5. Moussaif Intervention Court

يرفض action إذا فشل constraint/identifiability/sensitivity.

Tool 5/5 · unsafe selection rate

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-03

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Counterfactual Planner	state + candidates	counterfactual tree	branch coverage
2	Moussaif World-Effect Simulator	world model + action	effect distributions	forecast calibration
3	Moussaif Reversibility Evaluator	action contract	RevCost + checkpoints	rollback success
4	Moussaif Horizon Comparator	effect branches	horizon matrix	cross-horizon consistency
5	Moussaif Intervention Court	CEV packet	verdict	unsafe selection rate

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-03

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop؛ MCAI؛ M-RWEL؛ Authorization T0-T4؛ Project World Model | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Counterfactual Pair Training

تعلم تفضيل action من paired futures.

بيانات التدريب: same state, alternate actions

هدف التعلم: pairwise preference

مقياس النجاح: action ranking

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Outcome-Model Training

تدريب world-effect predictor مستقلا عن policy.

بيانات التدريب: state-action-nextstate

هدف التعلم: forecast loss

مقياس النجاح: Brier/ECE

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-03

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Reversible Intervention Curriculum

البدء بأفعال قابلة للعكس ثم توسيع blast radius.

Input: graded environments | **Output/Objective:** safety curriculum | **Metric:** rollback competence

4. Moussaif Delayed-Effect RL

إسناد reward على horizons متعددة مع persistence checks.

Input: long episodes | **Output/Objective:** multi-horizon return | **Metric:** delayed success

5. Moussaif Causal World-Model Distillation

تقطير causal effect estimates إلى planner خفيف.

Input: verified effect ledger | **Output/Objective:** distillation | **Metric:** CEV fidelity

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: retrieval / memory / skill / preference / adapter / continued-pretraining.

M-AFI-03

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

.interactive environments، action alternatives، hidden consequences، rollback traces، delayed outcomes، no-op controls

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence

world-effect / external test / reproducible held-out performance.

Non-evidence alone

عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- ارتفاع اختيار probes القابلة للعكس عند عدم اليقين.
- انخفاض side effects غير المقصودة.
- تحسن long-horizon success.
- تحسن calibration بين predicted وobserved effect.

M-AFI-03

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

يضاف CounterfactualPlanner بين MCAI وAuthorization Gate، ويرتبط WorldEffectVerifier لإغلاق prediction error.

عقد النقل

يمكن فصل world model عن policy model لتقليل self-confirmation واختباره عبر providers.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-03",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-03

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> بيئة stateful software أو browser مع آثار متأخرة، قارن direct plan vs counterfactual CEV under hidden tests.

المكذّب الرئيسي

> إذا لم يحسن counterfactual planning النجاح أو السلامة مقارنة بمخطط مباشر equal-budget، أو إذا كانت world models غير معايرة لدرجة تجعل ranking عشوائيا.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-03	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-03

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

1. model hallucinated futures
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. counterfactual explosion
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. risk underestimation
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. reversibility illusion
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. horizon myopia
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

الحالات المضادة للواقع لا تبرر actions محظورة؛ القيود الدستورية تعمل قبل CEV.

الحدود

counterfactual quality محدودة بجودة world model؛ لا ينبغي تضخيم precision الاصطناعي لصيغ CEV.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-03

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: World-Effect Loop، MCAI، M-RWEL، Authorization T0-T4، Project World Model | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للتخطيط المضاد للواقع لأثر العالم Moussaif Counterfactual World-Effect Planning Theory
الفارق الفريد	ينقل World-Effect من verifier بعدي إلى primitive تخطيطي قبلي مع counterfactuals متعددة الآفاق.
المعادلة	CEV(a)=Σ_h γ_h E[U(W_h^{do(a)})] - α Risk(a) - β RevCost(a) - χ Unc(a)
المكذّب	إذا لم يحسن counterfactual planning النجاح أو السلامة مقارنة بمخطط مباشر equal-budget، أو إذا كانت world models غير معايرة لدرجة تجعل ranking عشوائيا.
الأدوات 5	Moussaif Counterfactual Planner, Moussaif World-Effect Simulator, Moussaif Reversibility Evaluator, Moussaif Horizon Comparator, Moussaif Intervention Court
التدريب 5	Moussaif Counterfactual Pair Training, Moussaif Outcome-Model Training, Moussaif Reversible Intervention Curriculum, Moussaif Delayed-Effect RL, Moussaif Causal World-Model Distillation

التنبؤات القابلة للفحص

- ارتفاع اختيار probes القابلة للعكس عند عدم اليقين.
- انخفاض side effects غير المقصودة.
- تحسن long-horizon success.
- تحسن calibration بين predicted وobserved effect.

مثال تفسيري

قبل تعديل إعداد إنتاجي، يقارن النظام patch مباشر، no-op، canary، dry-run، وrollback simulation ثم يختار canary إذا حقق قيمة قريبة مع risk أقل.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-04 · MUT-AFI

نظرية مساعف للمزامنة الحاكمية متعددة الآفاق

Moussaif Multi-Horizon Governance Synchronization Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
GSync = Σ_h ω_h Align(G_h,E_h) - λ CrossHorizonConflict(E_1..E_H)
```

الفكرة المركزية: الحاكمية ليست gate لحظيا فقط؛ يجب مزامنة الهدف والقيود والأثر عبر آفاق زمنية متعددة، وكشف التعارض بين policy قصيرة المدى والسياسة العليا طويلة المدى قبل ترقية الفعل أو التعلم.

M-AFI-04

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

توجد حاكمية وWorld-Effect وحلقات محدودة، لكن long-horizon agency قد يحقق هدفا آنيا على حساب قيد أو قيمة تظهر بعد ساعات أو دورات لاحقة.

الأطروحة المقترحة

الحاكمية ليست gate لحظيا فقط؛ يجب مزامنة الهدف والقيود والأثر عبر آفاق زمنية متعددة، وكشف التعارض بين policy قصيرة المدى والسياسة العليا طويلة المدى قبل ترقية الفعل أو التعلم.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **يضيف temporal governance consistency ككمية مستقلة عن نجاح المهمة اللحظي.**

العلاقة بالسلالة السابقة

Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-04

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

> X = cross-horizon conflict، ω_h = horizon weight، E_h = predicted effect، G_h = governance constraints at horizon h

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-04",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "G_h = governance constraints at horizon h؛ E_h = predicted
effect؛ ω_h = horizon weight؛ X = cross-horizon conflict",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

> **حد معرفي:** المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-04

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
GSync = Σ_h ω_h Align(G_h,E_h) - λ CrossHorizonConflict(E_1..E_H)
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-04

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تقسيم المهمة إلى horizons تشغيلية وقيمية.

2. المرحلة
إسقاط constraints/goals على كل horizon.

3. المرحلة
توليد effect expectations لكل action/loop.

4. المرحلة
كشف conflict مثل short-term gain / long-term debt.

5. المرحلة
اختيار policy robust أو طلب approval إذا conflict جوهري.

6. المرحلة
تحديث horizon weights من آثار فعلية دون تعديل الدستور.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-04

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يُختبر على نماذج مختلفة مع نفس mission state لضمان أن sync ليس نمط لغة خاصا.

> **الإضافة غير القابلة للاختزال**
> يضيف temporal governance consistency ككمية مستقلة عن نجاح المهمة اللحظي.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

HorizonGovernanceMatrix يحيط Objective Hierarchy وCounterfactual Planner ويغذي M-TAOS scheduler.

M-AFI-04

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Horizon Governance Matrix
يمثل goals/constraints/effects عبر h.

Tool 1/5 coverage

2. Moussaif Temporal Conflict Detector
يكشف تعارض الأهداف/القيود بين الآفاق.

Tool 2/5 conflict recall

3. Moussaif Policy Synchronizer
يقترح policy مشتركة أو staggered plan.

Tool 3/5 GSync

4. Moussaif Delayed Consequence Ledger
يربط القرارات بآثار لاحقة فعلية.

Tool 4/5 attribution coverage

5. Moussaif Multi-Horizon Court
يحكم على short win / long loss وغيرها.

Tool 5/5 late-regression rate

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-04

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded، HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Horizon Governance Matrix	mission + policies	horizon matrix	coverage
2	Moussaif Temporal Conflict Detector	matrix	conflict graph	conflict recall
3	Moussaif Policy Synchronizer	conflicts + candidates	synchronized policy	GSync
4	Moussaif Delayed Consequence Ledger	events	delayed links	attribution coverage
5	Moussaif Multi-Horizon Court	policy packet	verdict	late-regression rate

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-04

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Multi-Horizon Return Training

تعلم منفعة موزعة زمنيا.

بيانات التدريب: long episodes

هدف التعلم: weighted return

مقياس النجاح: long-term success

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval,
regression}
```

2. Moussaif Temporal Constraint Curriculum

زيادة طول تأخر انتهاك القيود تدريجيا.

بيانات التدريب: delayed-violation tasks

هدف التعلم: constraint loss

مقياس النجاح: violation latency robustness

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval,
regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-04

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Long-Horizon Counterfactual Training

مقارنة بدائل لها آثار بعيدة مختلفة.

Input: counterfactual trajectories | **Output/Objective:** preference loss | **Metric:** horizon ranking

4. Moussaif Horizon-Conflict Preference Training

تفضيل الحل المتزن عند tradeoff موثق.

Input: policy pairs | **Output/Objective:** preference optimization | **Metric:** GSync preference

5. Moussaif Governance Synchronization Distillation

تقطير قرار المحكمة متعددة الآفاق إلى planner.

Input: court verdicts | **Output/Objective:** distillation | **Metric:** cheap sync fidelity

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-04

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

long-horizon tasks، delayed regressions، maintenance episodes، cost over time، policy drift traces.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- انخفاض goal debt المتراكم.
- تحسن الاستقرار بعد interruptions.
- خفض repeated short-term fixes.
- زيادة اختيار حلول أقل لمعانا لكنها أكثر استدامة.

M-AFI-04

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

HorizonGovernanceMatrix يحيط Objective Hierarchy وCounterfactual Planner ويغذي M-TAOS scheduler.

عقد النقل

يُختبر على نماذج مختلفة مع نفس mission state لضمان أن sync ليس نمط لغة خاصا.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-04",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-04

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> tasks من 50-200 خطوة مع delayed constraint violations؛ قارن agent لحظي مقابل synchronized governance.

المكذّب الرئيسي

> إذا لم يخفض GSync failures الطويلة أو يمكن استبداله بـdiscounted reward بسيط بنفس الأداء والسلامة.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-04	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-04

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

1. horizon blindness
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. over-discounting
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. policy oscillation
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. constraint dilution
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. future speculation
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

الأفق الطويل لا يسمح بتجاوز قيود آنية؛ الأعلى رتبة يظل نافذا في كل h.

الحدود

الأثر البعيد غير مؤكد؛ الأوزان الزمنية normative جزئيا ولا يمكن "اكتشافها" من البيانات وحدها.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-04

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + الأصول المتصلة: Governance Closure، Multi-Horizon World Effect، Objective Hierarchy، Bounded HyperLoops، M-TAOS | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للمزامنة الحاكمية متعددة الآفاق Moussaif Multi-Horizon Governance Synchronization Theory
الفارق الفريد	يضيف temporal governance consistency ككمية مستقلة عن نجاح المهمة اللحظي.
المعادلة	GSync = Σ_h ω_h Align(G_h,E_h) - λ CrossHorizonConflict(E_1..E_H)
المكذّب	إذا لم يخفض GSync failures الطويلة أو يمكن استبداله بـdiscounted reward بسيط بنفس الأداء والسلامة.
الأدوات 5	Moussaif Horizon Governance Matrix, Moussaif Temporal Conflict Detector, Moussaif Policy Synchronizer, Moussaif Delayed Consequence Ledger, Moussaif Multi-Horizon Court
التدريب 5	Moussaif Multi-Horizon Return Training, Moussaif Temporal Constraint Curriculum, Moussaif Long-Horizon Counterfactual Training, Moussaif Horizon-Conflict Preference Training, Moussaif Governance Synchronization Distillation

التنبؤات القابلة للفحص

- انخفاض goal debt المتراكم.
- تحسن الاستقرار بعد interruptions.
- خفض repeated short-term fixes.
- زيادة اختيار حلول أقل لمعانا لكنها أكثر استدامة.

مثال تفسيري

حل سريع يقلل latency الآن لكنه يخلق dependency هشاً بعد يوم. النظرية تضع long-horizon maintenance cost داخل الحكم قبل التنفيذ.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-05 · MUT-AFI

نظرية مساعف لإعادة البناء الحافظة للقيمة

Moussaif Value-Preserving Self-Reconstruction Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
L_total = L_rec + β L_val + η L_prov + ξ L_cap; Promote only if
invariant_gates = 1
```

الفكرة المركزية: أي إعادة بناء معرفية أو تشغيلية يجب أن تحقق invariants للهوية والقيم والسلطة والمصدر بالتوازي مع تحسين الأداء؛ التحسن الذي يخرق invariant يعد regression ولو رفع benchmark.

M-AFI-05

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

Self-Reconstruction Memory تحفظ lineage والتغير، لكن adaptation المستمر قد يعيد بناء schema/skills/memory بطريقة تحافظ على المعلومات وتفقد القيم أو الحدود الحاكمية.

الأطروحة المقترحة

أي إعادة بناء معرفية أو تشغيلية يجب أن تحقق invariants للهوية والقيم والسلطة والمصدر بالتوازي مع تحسين الأداء، التحسن الذي يخرق invariant يعد regression ولو رفع benchmark.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **يحول self-reconstruction من استعادة حالة إلى optimization مقيد يحمي القيم والنسب والحدود عبر إعادة التنظيم والتعلم.**

العلاقة بالسلالة السابقة

Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity AutoDeployment، Continuous Post-Training

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-05

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

L_rec = reconstruction loss؛ L_val = value/governance drift؛ L_prov = provenance loss؛ L_cap = capability regression

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-05",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "L_rec = reconstruction loss؛ L_val = value/governance drift؛
L_prov = provenance loss؛ L_cap = capability regression",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-05

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
L_total = L_rec + β L_val + η L_prov + ξ L_cap; Promote only if invariant_gates = 1
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-05

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تعريف invariant registry قبل أي reconstruction.

2. المرحلة
إعادة بناء candidate state من events/artifacts.

3. المرحلة
مقارنة candidate بالقيم والسلطة والمصدر والحساسية.

4. المرحلة
تشغيل capability regressions وnegative controls.

5. المرحلة
رفض أو عزل التغير المخالف ولو حسن metric محلي.

6. المرحلة
كتابة migration proof وrollback pointer.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-05

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

الـinvariants provider-neutral؛ يجب أن تبقى عند نقل state بين نماذج.

الإضافة غير القابلة للاختزال
يحول self-reconstruction من استعادة حالة إلى optimization مقيد يحمي القيم والنسب والحدود عبر إعادة التنظيم والتعلم.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

يعمل داخل ValueInvariantRegistry وMemory Promotion وSchema Migration وKnowledge-to-Kernel Delta Deployment.

M-AFI-05

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Value Invariant Registry
يحفظ الثوابت التي لا يجوز أن تضيع في reconstruction.

Tool 1/5 | invariant coverage

2. Moussaif Reconstruction Diff Analyzer
يقارن state قبل/بعد على semantics والسلطة والمصدر.

Tool 2/5 | critical-diff recall

3. Moussaif Provenance-Preserving Merger
يدمج knowledge/memory دون قطع lineage.

Tool 3/5 | lineage fidelity

4. Moussaif Value Drift Alarm
يرصد تغير policy implications بعد تحديث غير policy.

Tool 4/5 | false negative rate

5. Moussaif Reconstruction Court
يقرر promote/quarantine/rollback.

Tool 5/5 | post-promotion regression

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-05

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Value Invariant Registry	constitution/policies	versioned invariants	invariant coverage
2	Moussaif Reconstruction Diff Analyzer	state versions	typed diff	critical-diff recall
3	Moussaif Provenance-Preserving Merger	candidate patches	merged state + provenance	lineage fidelity
4	Moussaif Value Drift Alarm	before/after evals	drift alerts	false negative rate
5	Moussaif Reconstruction Court	diff + tests	verdict	post-promotion regression

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-05

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Invariant-Preserving Rewrite Training

إعادة كتابة/ضغط دون فقد constraints.

بيانات التدريب: paired full/compressed states

هدف التعلم: constraint preservation loss

مقياس النجاح: invariant retention

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Memory Reconstruction Pair Training

استعادة state من event subsets.

بيانات التدريب: event→state pairs

هدف التعلم: reconstruction loss

مقياس النجاح: state fidelity

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-05

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Rollback-Aware Adaptation

تعلم updates يمكن عكسها وتتبعها.

Input: versioned deltas | **Output/Objective:** reversible update objective | **Metric:** rollback success

4. Moussaif Provenance Restoration Training

استنتاج روابط المصدر المفقودة مع abstention عند الغموض.

Input: provenance graphs | **Output/Objective:** link prediction + abstain | **Metric:** provenance precision

5. Moussaif Value-Drift Adversarial Training

حقن تحسينات أداء تخفي policy drift.

Input: adversarial migrations | **Output/Objective:** drift detection loss | **Metric:** drift robustness

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-05

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

memory events، version histories، migrations، policy tests، corrupted patches، rollback traces.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- انخفاض silent value drift.
- تحسن rollback fidelity.
- انخفاض contamination من memory patches.
- وضوح أكبر بين capability update وpolicy update.

M-AFI-05

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

ValueInvariantRegistry يعمل داخل Memory Promotion وSchema Migration وKnowledge-to-Kernel Delta Deployment.

عقد النقل

الـinvariants provider-neutral؛ يجب أن تبقى عند نقل state بين نماذج.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-05",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-05

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> نفّذ schema/memory migrations مع injected goal/provenance corruption؛ قارن reconstruction عادي مقابل value-preserving.

المكذّب الرئيسي

> إذا لم تتنبأ invariant gates بانحرافات مستقبلية أو كانت مجرد تكرار للاختبارات التقليدية.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-05	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-05

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

1. **invariant overconstraint**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. **value freeze**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. **provenance loss**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. **migration poisoning**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. **hidden policy drift**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

القيم العليا ليست learnable تلقائيا من النجاح؛ تعديلها يحتاج authorization مستقل.

الحدود

تعريف values/invariants نفسه يحتاج حاكمية بشرية/مؤسسية ولا تحله الخوارزمية وحدها.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-05

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Self-Reconstruction Memory، Constitutional Precedence، Knowledge-Integrity، AutoDeployment، Continuous Post-Training | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لإعادة البناء الحافظة للقيمة Moussaif Value-Preserving Self-Reconstruction Theory
الفارق الفريد	يحول self-reconstruction من استعادة حالة إلى optimization مقيد يحمي القيم والنسب والحدود عبر إعادة التنظيم والتعلم.
المعادلة	L_total = L_rec + β L_val + η L_prov + ξ L_cap; Promote only if invariant_gates = 1
المكذّب	إذا لم تتنبأ invariant gates بانحرافات مستقبلية أو كانت مجرد تكرار للاختبارات التقليدية.
الأدوات 5	Moussaif Value Invariant Registry, Moussaif Reconstruction Diff Analyzer, Moussaif Provenance-Preserving Merger, Moussaif Value Drift Alarm, Moussaif Reconstruction Court
التدريب 5	,Moussaif Invariant-Preserving Rewrite Training, Moussaif Memory Reconstruction Pair Training, Moussaif Rollback-Aware Adaptation Moussaif Provenance Restoration Training, Moussaif Value-Drift Adversarial Training

التنبؤات القابلة للفحص

- انخفاض silent value drift.
- تحسن rollback fidelity.
- انخفاض contamination من memory patches.
- وضوح أكبر بين capability update وpolicy update.

مثال تفسيري

إذا ضغط النظام ذاكرة مهمة واحتفظ بالنتيجة لكنه حذف أن الفعل كان يتطلب approval، فإعادة البناء فاشلة رغم سلامة المحتوى الظاهري.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-06 · MUT-AFI

نظرية مساعف للاستتباب الدلالي عبر النماذج

Moussaif Cross-Model Semantic Homeostasis Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
H_sem = 1 - D_sem(S*, Normalize(S_j)); TriggerCorrection if H_sem < τ_h or dH_sem/dn < -δ
```

الفكرة المركزية: يجب أن يمتلك النظام مرجعاً دلالياً معيارياً وآلية تغذية راجعة تقيس الانزياح بعد كل عبور بين نموذج أو وكيل أو ذاكرة، ثم تعيد الحالة إلى نطاق دلالي مقبول قبل السماح بالتراكم.

M-AFI-06

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

تغطي M-RST ترجمة حالة الاستدلال بين مزودين، لكن التبديل المتكرر والتلخيص وإعادة الحقن قد يراكم انزياحاً دلالياً لا يظهر في اختبار نقل واحد.

الأطروحة المقترحة

يجب أن يمتلك النظام مرجعاً دلالياً معيارياً وآلية تغذية راجعة تقيس الانزياح بعد كل عبور بين نموذج أو وكيل أو ذاكرة، ثم تعيد الحالة إلى نطاق دلالي مقبول قبل السماح بالتراكم.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **المتغير الجديد هو الاستقرار الدلالي عبر سلسلة من التحويلات، لا جودة ترجمة منفردة. النظرية تقيس انجراف المعنى الذي يغير قراراً أو دليلاً أو قيداً، لا اختلاف الصياغة السطحي.**

العلاقة بالسلالة السابقة

M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-06

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

S*: الحالة الدلالية المرجعية؛ S_j: إسقاط المضيف j؛ D_sem: مسافة دلالية مقيدة بالقرار؛ H_sem: درجة الاستتباب؛ n: عدد عمليات العبور.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-06",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "S*: الحالة الدلالية المرجعية؛ S_j: إسقاط المضيف j؛ D_sem: مسافة
دلالية مقيدة بالقرار؛ H_sem: درجة الاستتباب؛ n: عدد عمليات العبور.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-06

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
H_sem = 1 - D_sem(S*, Normalize(S_j)); TriggerCorrection if H_sem < τ_h or dH_sem/dn < -δ
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-06

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
إنشاء Semantic Checksum من الهدف والادعاءات والأدلة والقرارات والمخاطر والأسئلة المفتوحة.

2. المرحلة
ترجمة الحالة إلى المضيف الجديد مع فصل البنية عن الألفاظ.

3. المرحلة
إسقاط الناتج إلى Canonical State وإعادة حساب D_sem على الحقول ذات الأثر.

4. المرحلة
تشغيل اختبار قرار مضاد: هل ينتج المضيف قراراً مختلفاً بسبب الترجمة لا بسبب دليل جديد؟

5. المرحلة
إعادة ترطيب الحقول المنحرفة من المصدر المعياري أو ledger عند تجاوز العتبة.

6. المرحلة
تعلم توقيع انزياح خاص بكل مزود/سلسلة تحويل لتحسين Normalizer مستقبلاً.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-06

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

ترتبط بـM-RST للنقل، وROFSC للضغط، وSelf-Reconstruction لإعادة البناء، لكنها تضيف ديناميكا زمنية للاستقرار.

الإضافة غير القابلة للاختزال
المتغير الجديد هو الاستقرار الدلالي عبر سلسلة من التحويلات، لا جودة ترجمة منفردة. النظرية تقيس انجراف المعنى الذي يغير قراراً أو دليلاً أو قيداً، لا اختلاف الصياغة السطحي.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

طبقة مراقبة بين M-RST وMemory Writeback؛ تعمل عند model swap، compaction، handoff، أو إعادة استئناف مهمة.

M-AFI-06

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Semantic State Canonicalizer
يحوّل حالات متعددة الصياغة إلى مخطط معياري قابل للمقارنة.

Tool 1/5 فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Provider Translation Matrix
يسجل أخطاء التحويل واتجاهاتها بين الأزواج المختلفة من النماذج.

Tool 2/5 فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Drift Sentinel
يراقب مشتقة الانزياح عبر الزمن ويطلق gate عند تجاوز العتبة.

Tool 3/5 فائدة قابلة للقياس تحت baseline وablation

4. Moussaif State Capsule Comparator
يقارن القرارات والقيود والأدلة قبل/بعد النقل لا الكلمات فقط.

Tool 4/5 فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Semantic Homeostasis Court
يحكم هل الانحراف تعلم مشروع أم drift يحتاج استعادة.

Tool 5/5 فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-06

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Semantic State Canonicalizer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Provider Translation Matrix	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Drift Sentinel	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif State Capsule Comparator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Semantic Homeostasis Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-06

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Cross-Provider Paraphrase Invariance Training

تدريب على حفظ القيود والادعاءات رغم اختلاف الأسلوب بين مزودين.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Name-Stripped State Transfer Training

إزالة أسماء النظريات واختبار نقل الوظيفة نفسها.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-06

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Provider-Swap Curriculum

زيادة عدد تبديلات المضيف تدريجياً مع مهام طويلة.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Drift-Correction Self-Supervision

توليد أزواج حالة منجرفة/مصححة من traces حقيقية.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Semantic Checksum Training

تعليم النموذج استخراج invariants الحاكمية إلى سجل صغير قابل للتحقق.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-06

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

سجلات canonical states، تبديلات مزودين، source pointers، decisions، constraints، وأخطاء اتساق بعد النقل.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- الانزياح يتراكم لا خطياً مع عدد التحويلات غير المصححة.
- الحقول الحاكمية تنجرف أقل عند حمايتها كـchecksum نمطي.
- التغذية الراجعة تقلل الانحراف حتى مع نماذج مختلفة.
- قد تزيد الكلفة قليلاً لكنها تخفض أخطاء الاستمرارية عالية الأثر.

M-AFI-06

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

طبقة مراقبة بين M-RST وMemory Writeback؛ تعمل عند model swap، compaction، handoff، أو إعادة استئناف مهمة.

عقد النقل

ترتبط بـM-RST للنقل، وROFSC للضغط، وSelf-Reconstruction لإعادة البناء، لكنها تضيف ديناميكا زمنية للاستقرار.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-06",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-06

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مسار طويل من 20 تبديل Provider/Agent مع مهام تتطلب الحفاظ على قيود خفية غير سرية. قارن no-normalization وone-shot RST وhomeostatic feedback على Decision Fidelity وConstraint Retention.

المكذّب الرئيسي

> إذا لم يزد الاستتباب الدلالي دقة القرار أو الاستمرارية تحت تبديل المزودين مقارنة بتمرير StateCapsule بسيط متساوي الميزانية، تضعف النظرية.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-06	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-06

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed MUT-AFI synthesis

1. تقديس الصياغة المرجعية بدلاً من المعنى.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. مقياس تشابه لا يلتقط تغير القرار.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. تصحيح مفرط يمحو تحسيناً مشروعاً.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. تلوث S* بمعلومة قديمة.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. حلقة إعادة ترطيب غير محدودة.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا تحفظ سلسلة التفكير الخفية؛ تحفظ فقط الحالة القابلة للتصريح. التغيير المتعمد والمأذون في الهدف يجب أن يحدث عبر Goal Update وليس عبر مصحح الانزياح.

الحدود

تعريف D_sem يعتمد على المهمة؛ لا يوجد مقياس واحد عالمي للمعنى. الاستتباب لا يعني ثباتاً مطلقاً بل ثبات invariants مع السماح بالتعلم.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-06

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory | الحالة: Proposed

الهوية	نظرية مساعف للاستتباب الدلالي عبر النماذج Moussaif Cross-Model Semantic Homeostasis Theory
الفارق الفريد	المتغير الجديد هو الاستقرار الدلالي عبر سلسلة من التحويلات، لا جودة ترجمة منفردة. النظرية تقيس انجراف المعنى الذي يغير قراراً أو دليلاً أو قيداً، لا اختلاف الصياغة السطحي.
المعادلة	H_sem = 1 - D_sem(S*, Normalize(S_j)); TriggerCorrection if H_sem < τ_h or dH_sem/dn < -δ
المكذّب	إذا لم يزد الاستتباب الدلالي دقة القرار أو الاستمرارية تحت تبديل المزودين مقارنة بتمرير StateCapsule بسيط متساوي الميزانية، تضعف النظرية.
الأدوات 5	Moussaif Semantic State Canonicalizer, Moussaif Provider Translation Matrix, Moussaif Drift Sentinel, Moussaif State Capsule Comparator, Moussaif Semantic Homeostasis Court
التدريب 5	Moussaif Cross-Provider Paraphrase Invariance Training, Moussaif Name-Stripped State Transfer Training, Moussaif Provider-Swap Curriculum, Moussaif Drift-Correction Self-Supervision, Moussaif Semantic Checksum Training

التنبؤات القابلة للفحص

- الانزياح يتراكم لا خطياً مع عدد التحويلات غير المصححة.
- الحقول الحاكمية تنجرف أقل عند حمايتها كـchecksum نمطي.
- التغذية الراجعة تقلل الانحراف حتى مع نماذج مختلفة.
- قد تزيد الكلفة قليلاً لكنها تخفض أخطاء الاستمرارية عالية الأثر.

مثال تفسيري

وكيل يختصر شرطاً «لا نشر خارجي» ثم يمرره لثلاثة نماذج. التشابه النصي قد يبقى مرتفعاً بينما يختفي قيد السلطة. Semantic checksum يكتشف غياب الحقل الحاكم ويوقف التنفيذ.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-07 · MUT-AFI

نظرية مساعف لإغلاق القدرة التركيبية

Moussaif Compositional Capability Closure Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Closure = (Π_j g_j) × min_j(r_j) × max(0, 1 + Σ_{i<j} I_ij)
```

الفكرة المركزية: القدرة العامة في نظام وكيلي خاصية لشبكة تركيب مغلقة؛ لا تُعد القدرة مكتملة إلا إذا كان كل رابط لازم بين الإدراك والتخطيط والتنفيذ والتحقق والذاكرة صالحاً وقابلاً للاختبار.

M-AFI-07

1. الفجوة البحثية والأطروحة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
توجد مهارات وقدرات وحلقات وأدوات كثيرة، لكن وجود المكونات لا يثبت أن تركيبها يغلق مهمة كاملة من البداية إلى أثر عالم متحقق.

الأطروحة المقترحة

القدرة العامة في نظام وكيلي خاصية لشبكة تركيب مغلقة؛ لا تُعد القدرة مكتملة إلا إذا كان كل رابط لازم بين الإدراك والتخطيط والتنفيذ والتحقق والذاكرة صالحاً وقابلاً للاختبار.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تنتقل وحدة القياس من عدد المهارات إلى قابلية إغلاق مهمة مركبة. عنق زجاجة واحد قد يجعل ألف مهارة غير كافية.**

العلاقة بالسلالة السابقة
HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-07

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

g_j بوابات سلامة المكون؛ r_j جاهزية الرابط؛ I_ij تآثر موجب/سالب بين مكونين؛ C شبكة الإغلاق.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-07",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "g_j بوابات سلامة المكون؛ r_j جاهزية الرابط؛ I_ij تآثر موجب/سالب بين مكونين؛ C شبكة الإغلاق.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-07

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Closure = (Π_j g_j) × min_j(r_j) × max(0, 1 + Σ_{i<j} I_ij)
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-07

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تحويل المهمة إلى Capability Dependency Graph يربط /perception/retrieval/reasoning/action/verification memory.

2. المرحلة
إسناد readiness لكل عقدة ولكل edge بدليل اختبار لا بالاسم.

3. المرحلة
تشغيل component-dropout لمعرفة الروابط الضرورية فعلاً.

4. المرحلة
قياس تآثرات synergy/interference بين skill وtool وloop وagent.

5. المرحلة
اكتشاف minimum cut الذي يمنع World-Effect closure.

6. المرحلة
ترقية القدرة فقط بعد نجاح composition على أكثر من عائلة مهمة.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-07

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

تكمّل HyperAbility بإضافة معيار اكتمال بين القدرات، وتكمّل HAKIM-MCD بقياس صلاحية السلسلة كاملة.

> **الإضافة غير القابلة للاختزال**
> تنتقل وحدة القياس من عدد المهارات إلى قابلية إغلاق مهمة مركبة. عنق زجاجة واحد قد يجعل ألف مهارة غير كافية.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

> **قاعدة MUT-AFI:** هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Compiler يبني graph من Task Contract ثم Router يختار أقل subgraph يغلق تعريف الإنجاز.

M-AFI-07

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Capability Closure Graph

يبني DAG للقدرات والروابط الضرورية.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Missing-Link Diagnoser

يحدد edge أو contract المفقود الذي يمنع الإغلاق.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Composition Harness

يشغل تراكيب محددة تحت نفس الميزانية ويقارنها.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Bottleneck Profiler

يحسب minimum cuts وعنق الزجاجة السببي.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Closure Court

يرفض ادعاء القدرة إذا لم تُغلق المهمة حتى World Effect.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-07

7. عقود الأدوات ومقاييسها

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Capability Closure Graph	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Missing-Link Diagnoser	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Composition Harness	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Bottleneck Profiler	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Closure Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-07

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Component-Dropout Training

إخفاء مكونات عشوائياً لتعلم التعويض وحدود الاعتماد.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Compositional Curriculum

تدرج من قدرات ثنائية إلى سلاسل طويلة ذات تحقق.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-07

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Skill-Loop-Tool Alignment Training

تعليم العقود الصحيحة بين طبقات التنفيذ.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Closure-Reward Training

مكافأة completion verified بدلاً من نجاح خطوات محلية.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Cross-Domain Composition Transfer

اختبار نفس بنية التركيب في مجالات جديدة بأسماء مختلفة.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-07

10. البيانات والتقييم والمعايير

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

نتائج مكونات منفردة، traces تركيبية، effect verification، failures، dependencies، وablation logs.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- نجاح المكونات منفردة يبالغ في تقدير الأداء التركيبي.
- minimum-cut صغير يفسر نسبة كبيرة من الفشل.
- التدريب على الروابط أحياناً أجدى من تدريب مكون جديد.
- closure يزداد بشكل قفزي عند إصلاح bottleneck حاكم.

M-AFI-07

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Compiler يبني graph من Task Contract ثم Router يختار أقل subgraph يغلق تعريف الإنجاز.

عقد النقل

تكّمل HyperAbility بإضافة معيار اكتمال بين القدرات، وتكّمل HAKIM-MCD بقياس صلاحية السلسلة كاملة.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-07",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-07

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مجموعة مهام متعددة الخطوات، قارن اختيار المكونات حسب أعلى scores فردية مقابل closure-aware composition مع ميزانية متساوية.

المكذّب الرئيسي

> إذا تنبأ مجموع scores الفردية بالنجاح مثل Closure Graph أو أفضل منه تحت مهام مركبة جديدة، لا تضيف النظرية قيمة.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-07	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-07

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

1. تحويل كل dependency إلى إلزام ثابت.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. إهمال التداخل السلبي بين الأدوات.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. تقييم داخل نفس corpus فقط.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. إغلاق شكلي بلا World Effect.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. مضاعفة مكونات redundant لرفع score.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا يسمح closure score بتجاوز authorization أو evidence gates؛ بوابات g_j تشملها كعوامل صفرية.

الحدود

تقدير I_{ij} مكلف توافقيّاً؛ يلزم sparse interaction search. لا يثبت closure مجالاً عاماً خارج distribution المختبر.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-07

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لإغلاق القدرة التركيبية Moussaif Compositional Capability Closure Theory
الفارق الفريد	تنتقل وحدة القياس من عدد المهارات إلى قابلية إغلاق مهمة مركبة. عنق زجاجة واحد قد يجعل ألف مهارة غير كافية.
المعادلة	Closure = (Π_j g_j) × min_j(r_j) × max(0, 1 + Σ_{i<j} I_ij)
المكذّب	إذا تنبأ مجموع scores الفردية بالنجاح مثل Closure Graph أو أفضل منه تحت مهام مركبة جديدة، لا تضيف النظرية قيمة.
الأدوات 5	Moussaif Capability Closure Graph, Moussaif Missing-Link Diagnoser, Moussaif Composition Harness, Moussaif Bottleneck Profiler, Moussaif Closure Court
التدريب 5	Moussaif Component-Dropout Training, Moussaif Compositional Curriculum, Moussaif Skill-Loop-Tool Alignment Training, Moussaif Closure-Reward Training, Moussaif Cross-Domain Composition Transfer

التنبؤات القابلة للفحص

- نجاح المكونات منفردة يبالغ في تقدير الأداء التركيبي.
- minimum-cut صغير يفسر نسبة كبيرة من الفشل.
- التدريب على الروابط أحياناً أجدى من تدريب مكون جديد.
- closure يزداد بشكل قفزي عند إصلاح bottleneck حاكم.

مثال تفسيري

قد يملك النظام Skill للبحث وSkill للبرمجة وTool للنشر، لكنه يفتقد تحقق post-deploy؛ Closure=0 أو منخفض حتى يضاف رابط verification.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-08 · MUT-AFI

نظرية مساعف لتعلّم طوبولوجيا الفشل

Moussaif Failure Topology Learning Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
FailureRisk(v_j)=Σ_i p_ij·P(v_i|trace)·S_j ;
RepairValue(R_k)=ΔRisk×TransferScope/(ε+Cost)
```

الفكرة المركزية: للفشل طوبولوجيا سببية قابلة للتعلم: عقد تمثل signatures وروابط تمثل احتمال الانتقال أو الاشتراك في السبب. إصلاح البنية المشتركة يجب أن يتقدم على ترقيع الحالات المنفردة.

M-AFI-08

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
المحاكم تسجل الفشل، لكن قائمة أخطاء منفصلة لا تكشف البنية المشتركة أو مسارات الانتقال التي تجعل فشلاً محلياً يتكرر في مجالات مختلفة.

الأطروحة المقترحة

للفشل طوبولوجيا سببية قابلة للتعلم: عقد تمثل signatures وروابط تمثل احتمال الانتقال أو الاشتراك في السبب. إصلاح البنية المشتركة يجب أن يتقدم على ترقيع الحالات المنفردة.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تحويل سجل الفشل إلى graph ديناميكي يسمح بتوقع الفشل اللاحق ونقل الإصلاح عبر مهام مختلفة.**

العلاقة بالسلالة السابقة
Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-08

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

v_i عقد فشل؛ p_ij احتمال انتقال/سبب مشترك؛ S_j شدة؛ R_k إصلاح؛ TransferScope مجال النقل.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-08",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "v_i عقد فشل؛ p_ij احتمال انتقال/سبب مشترك؛ S_j شدة؛ R_k إصلاح؛
TransferScope مجال النقل.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-08

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
FailureRisk(v_j)=Σ_i p_ij·P(v_i|trace)·S_j ; RepairValue(R_k)=ΔRisk×TransferScope/(ε+Cost)
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-08

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تطبيع failure signatures من courts وtool errors وworld-effect mismatches.

2. المرحلة
بناء graph يربط co-occurrence والتسلسل والسبب المشترك المحتمل.

3. المرحلة
تجميع failures إلى motifs مثل stale-state→wrong-plan→unsafe-action.

4. المرحلة
تقدير transferability لإصلاح واحد عبر عقد متعددة.

5. المرحلة
التنبؤ بالعقد التالية عالية الاحتمال قبل وصولها.

6. المرحلة
تحديث graph بعد كل repair وقياس هل اختفى motif أم انتقل.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-08

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يوسع Failure Memory من سجل خطي إلى بنية قابلة للتنبؤ؛ يتكامل مع MCD عبر رفع risk/centrality.

الإضافة غير القابلة للاختزال
تحويل سجل الفشل إلى graph ديناميكي يسمح بتوقع الفشل اللاحق ونقل الإصلاح عبر مهام مختلفة.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

يعمل بعد Court verdict وقبل Evolution promotion؛ يمكنه أيضاً تغذية Router بقائمة risk motifs.

M-AFI-08

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً؛ كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Failure Graph Builder
يبني graph موحداً من signatures وآثار التنفيذ.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Failure Clusterer
يستخرج motifs متكررة عبر مجالات مختلفة.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Repair Transfer Engine
يقترح إصلاحاً عالي العائد على عدة عقد.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Failure Forecaster
يتنبأ بأقرب حالات الفشل من الحالة الحالية.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Failure Topology Court
يفصل الارتباط عن السبب ويطلب ablation عند الحاجة.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-08

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Failure Graph Builder	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Failure Clusterer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Repair Transfer Engine	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Failure Forecaster	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Failure Topology Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-08

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Negative-Trajectory Training

تدريب على مسارات فشل كاملة لا أمثلة خاطئة منفردة.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Failure Contrastive Training

تمييز signatures متشابهة سطحياً مختلفة سببياً.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-08

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Repair-Pair Training

ربط failure motif بإصلاح ونتيجة تحقق.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Recurrence-Prediction Training

توقع العقد التالية في graph من trace جزئي.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Adversarial Failure Synthesis

توليد حالات مركبة لاختبار robustness دون اعتمادها كدليل حقيقي.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-08

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

.failure traces، court verdicts، tool errors، rollback، task context، repair actions، outcomes

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- الفشل يتجمع في motifs أقل عدداً من الحالات الخام.
- إصلاح عقد مركزية يخفض فشلاً في مجالات لم تدرب عليها مباشرة.
- بعض الإصلاحات تغير شكل الفشل لا معدله.
- التكرار الزمني يكشف dependencies لا تظهر في taxonomy ثابتة.

M-AFI-08

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

يعمل بعد Court verdict وقبل Evolution promotion؛ يمكنه أيضاً تغذية Router بقائمة risk motifs.

عقد النقل

يوسع Failure Memory من سجل خطي إلى بنية قابلة للتنبؤ؛ يتكامل مع MCD عبر رفع risk/centrality.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-08",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-08

12. التجربة الحاسمة والـAblation

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> قسم traces زمنياً، تعلم topology على الماضي، وتنبأ بفشل المستقبل وقارن precision/recall وrepair transfer.

المكذّب الرئيسي

> إذا لم تتفوق تنبؤات graph وإصلاحاته على frequency-based failure prioritization في holdout زمني، تضعف الفرضية.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-08	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-08

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 Parts I-IV | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

1. اعتبار الارتباط سببية.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. دمج أخطاء مختلفة تحت label عام.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. نسيان نجاحات تمنع bias.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. Graph كثيف بلا utility.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. تعلم من synthetic failures فقط.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

أي توقع عالي الخطر يجب أن يرفع verification لا أن يحظر آلياً بلا evidence. يحتفظ graph بدرجة الثقة في edge.

الحدود

تحديد السبب يحتاج تدخلات/ablation، topology من الملاحظة وحدها قد تكون مخادعة.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-08

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Demolition Court، Failure Memory، Self-Reconstruction، Evolution Laboratory | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لتعلّم طوبولوجيا الفشل Moussaif Failure Topology Learning Theory
الفارق الفريد	تحويل سجل الفشل إلى graph ديناميكي يسمح بتوقع الفشل اللاحق ونقل الإصلاح عبر مهام مختلفة.
المعادلة	FailureRisk(v_j)=Σ_i p_ij·P(v_i|trace)·S_j ; RepairValue(R_k)=ΔRisk×TransferScope/(ε+Cost)
المكذّب	إذا لم تتفوق تنبؤات graph وإصلاحاته على frequency-based failure prioritization في holdout زمني، تضعف الفرضية.
الأدوات 5	Moussaif Failure Graph Builder, Moussaif Failure Clusterer, Moussaif Repair Transfer Engine, Moussaif Failure Forecaster, Moussaif Failure Topology Court
التدريب 5	-Moussaif Negative-Trajectory Training, Moussaif Failure Contrastive Training, Moussaif Repair-Pair Training, Moussaif Recurrence Prediction Training, Moussaif Adversarial Failure Synthesis

التنبؤات القابلة للفحص

- الفشل يتجمع في motifs أقل عدداً من الحالات الخام.
- إصلاح عقد مركزية يخفض فشلاً في مجالات لم تدرب عليها مباشرة.
- بعض الإصلاحات تغير شكل الفشل لا معدله.
- التكرار الزمني يكشف dependencies لا تظهر في taxonomy ثابتة.

مثال تفسيري

> أخطاء نشر خاطئ قد تبدو متنوعة، لكن graph يكشف motif: stale environment pointer → wrong target → verification skipped.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

> **Next Gate:** بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-09 · MUT-AFI

نظرية مساعف لتحويل عدم اليقين إلى وكالة

Moussaif Uncertainty-to-Agency Conversion Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
AgencyScore(a)=EV(a)+EIG(a)-Risk(a)·Irev(a)·U_rel(a);
a∈{ACT,QUERY,RETRIEVE,SIMULATE,ABSTAIN}
```

الفكرة المركزية: عدم اليقين ليس سبباً دائماً لمزيد من التفكير. يجب تفكيكه إلى عدم يقين قابل للإزالة وغير قابل للإزالة وربطه بقابلية الرجوع وقيمة المعلومات لتحديد أفضل فعل معرفي أو خارجي.

M-AFI-09

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
المعمارية تقيس uncertainty وتختار agency، لكن ينقص قانون يحدد متى يتحول عدم اليقين إلى سؤال أو استرجاع أو محاكاة أو فعل أو امتناع.

الأطروحة المقترحة

عدم اليقين ليس سبباً دائماً لمزيد من التفكير. يجب تفكيكه إلى عدم يقين قابل للإزالة وغير قابل للإزالة وربطه بقابلية الرجوع وقيمة المعلومات لتحديد أفضل فعل معرفي أو خارجي.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تحويل uncertainty من وصف إلى متغير قرار متعدد الأنماط، مع اعتبار الامتناع والسؤال أفعالاً معرفية لها قيمة وكلفة.**

العلاقة بالسلالة السابقة
MCAI، QMCD، Authorization، World-Effect

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-09

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

EV(a) قيمة الفعل، EIG(a) كسب المعلومات، Risk(a)، Irev(a) عدم الرجوع، U_rel عدم اليقين المتعلق بالقرار.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-09",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "EV(a) قيمة الفعل؛ EIG(a) كسب المعلومات؛ Risk(a)؛ Irev(a) عدم
الرجوع؛ U_rel عدم اليقين المتعلق بالقرار.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-09

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
AgencyScore(a)=EV(a)+EIG(a)-Risk(a)·Irev(a)·U_rel(a);
a∈{ACT,QUERY,RETRIEVE,SIMULATE,ABSTAIN}
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-09

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تفكيك عدم اليقين إلى epistemic/aleatoric/state/authority/tool uncertainty.

2. المرحلة
تحديد أي جزء يمكن خفضه بسؤال أو مصدر أو simulation.

3. المرحلة
تقدير reversibility وblast radius لكل فعل خارجي.

4. المرحلة
حساب EIG وEV تحت ميزانية محدودة.

5. المرحلة
اختيار mode وكالة لا action محتوى فقط.

6. المرحلة
إعادة القياس بعد كل observation لمنع loops بحثية غير منتجة.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-09

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يصل QMCD بجودة السؤال إلى MCAI/Agency، ويجعل World Effect جزءاً من قرار مقدار عدم اليقين المقبول.

الإضافة غير القابلة للاختزال
تحويل uncertainty من وصف إلى متغير قرار متعدد الأنماط، مع اعتبار الامتناع والسؤال أفعالاً معرفية لها قيمة وكلفة.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Agency mode gate بين MCAI candidate selection وExecutor.

M-AFI-09

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً؛ كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Uncertainty Decomposer

يفصل أنواع عدم اليقين ذات سياسات مختلفة.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Action-Query Gate

يختار بين الفعل والسؤال والاسترجاع والمحاكاة.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Reversibility-Aware Selector

يعدل القرار حسب تكلفة الرجوع ونطاق الأثر.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Abstention Planner

يحول الامتناع إلى قرار مفسر مع next-best action.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Uncertainty-to-Agency Court

يفحص هل الوكالة متناسبة مع uncertainty والسلطة.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-09

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Uncertainty Decomposer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Action-Query Gate	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Reversibility-Aware Selector	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Abstention Planner	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Uncertainty-to-Agency Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

موحّد Tool Contract

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-09

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Query-vs-Act Preference Training

أزواج تفضيل تعلم متى السؤال أفضل من الفعل.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Calibrated Abstention Training

تعليم الامتناع في الحالات التي لا تكفي الأدلة.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-09

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Uncertainty Decomposition Training

تصنيف مصدر uncertainty قبل القرار.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Reversible-Action Curriculum

البدء بأفعال قابلة للرجوع ثم رفع المخاطر تدريجياً.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Information-Gain RL

مكافأة المعلومة التي تغير قراراً فعلياً لا الاسترجاع الكثيف.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-09

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

approvals، query/retrieval gains، outcomes، reversibility، actions، uncertainty annotations، task states.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- ارتفاع uncertainty مع فعل قابل للرجوع قد يبرر ACT+verify.
- عدم يقين في authority يفرض QUERY/APPROVAL حتى لو المعرفة عالية.
- EIG المنخفض يوقف البحث الطويل.
- ABSTAIN الصحيح يحسن calibration لا يعد فشلاً.

M-AFI-09

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Agency mode gate بين MCAI candidate selection وExecutor.

عقد النقل

يصل QMCD بجودة السؤال إلى MCAI/Agency، ويجعل World Effect جزءاً من قرار مقدار عدم اليقين المقبول.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-09",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-09

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مهام مختلطة حيث بعضها يحتاج سؤالاً وبعضها يحتاج فعلاً سريعاً؛ قارن fixed policy وconfidence threshold وAgencyScore.

المكذّب الرئيسي

> إذا لم يحسن selector معدل النجاح/السلامة مقابل threshold بسيط على confidence مع نفس الميزانية، تُرفض قيمة التعقيد.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-09	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-09

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

1. خلط كل uncertainty في رقم واحد.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. مبالغة في كلفة عدم الرجوع.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. استخدام confidence ذاتي غير معاير.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. سؤال المستخدم فيما يمكن التحقق منه آلياً.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. الامتناع المزمن.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

Authority uncertainty يملك أسبقية مستقلة؛ لا يعوضه EV مرتفع. الأفعال غير القابلة للرجوع تحتاج gate أعلى.

الحدود

EIG وEV تقديرات تقريبية؛ يجب معايرتها بالمجال. لا يحل uncertainty غير المرصود.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-09

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: MCAI، QMCD، Authorization، World-Effect | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لتحويل عدم اليقين إلى وكالة Moussaif Uncertainty-to-Agency Conversion Theory
الفارق الفريد	تحويل uncertainty من وصف إلى متغير قرار متعدد الأنماط، مع اعتبار الامتناع والسؤال أفعالاً معرفية لها قيمة وكلفة.
المعادلة	AgencyScore(a)=EV(a)+EIG(a)-Risk(a)·Irev(a)·U_rel(a); a∈{ACT,QUERY,RETRIEVE,SIMULATE,ABSTAIN}
المكذّب	إذا لم يحسن selector معدل النجاح/السلامة مقابل threshold بسيط على confidence مع نفس الميزانية، تُرفض قيمة التعقيد.
الأدوات 5	Moussaif Uncertainty Decomposer, Moussaif Action-Query Gate, Moussaif Reversibility-Aware Selector, Moussaif Abstention Planner, Moussaif Uncertainty-to-Agency Court
التدريب 5	Moussaif Query-vs-Act Preference Training, Moussaif Calibrated Abstention Training, Moussaif Uncertainty Decomposition Training, Moussaif Reversible-Action Curriculum, Moussaif Information-Gain RL

التنبؤات القابلة للفحص

- ارتفاع uncertainty مع فعل قابل للرجوع قد يبرر ACT+verify.
- عدم يقين في authority يفرض QUERY/APPROVAL حتى لو المعرفة عالية.
- EIG المنخفض يوقف البحث الطويل.
- ABSTAIN الصحيح يحسن calibration لا يعد فشلاً.

مثال تفسيري

> تعديل ملف محلي reversible مع uncertainty متوسط قد ينفذ مع checkpoint، بينما إرسال بريد خارجي بنفس uncertainty يتطلب approval.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

> **Next Gate:** بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-10 · MUT-AFI

نظرية مساعف لموازنة الحمل الإبستيمي

Moussaif Epistemic Load Balancing Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Load_i = Debt_i·Cent_i·Risk_i·FreshPenalty_i/(ε+VCap_i); schedule verification to minimize max_i Load_i
```

الفكرة المركزية: يجب إدارة المعرفة كشبكة تحمل ديناً تحققياً. تتلقى العقد ذات الدين والمركزية والخطر والقدم Verification Budget أكبر حتى لا تصبح نقطة فشل نظامية.

M-AFI-10

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

Sparse routing يوزع compute، لكن تراكم claims غير المتحققة أو القديمة أو المركزية يخلق حملاً معرفياً قد يختبئ حتى ينهار قرار لاحق.

الأطروحة المقترحة

يجب إدارة المعرفة كشبكة تحمل ديناً تحققياً. تتلقى العقد ذات الدين والمركزية والخطر والقدم Verification Budget أكبر حتى لا تصبح نقطة فشل نظامية.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **يميز بين compute load وepistemic load. عقدة صغيرة نصياً قد تحمل عبئاً كبيراً إذا اعتمدت عليها قرارات كثيرة.**

العلاقة بالسلالة السابقة

Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-10

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

Debt_i دين تحقق؛ Cent_i مركزية؛ Risk_i؛ FreshPenalty_i؛ VCap_i سعة التحقق.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-10",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "Debt_i ؛دين تحقق Cent_i ؛مركزية Risk_i؛ FreshPenalty_i؛ VCap_i
سعة التحقق.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-10

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Load_i = Debt_i·Cent_i·Risk_i·FreshPenalty_i/(ε+VCap_i); schedule verification to minimize max_i Load_i
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-10

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
حساب debt لكل claim من عمره ونوع المصدر وحالة التحقق.

2. المرحلة
ضرب الدين في centrality وrisk لإظهار النقاط النظامية.

3. المرحلة
تقدير Verification Capacity المتاحة من أدوات/وقت/مصادر.

4. المرحلة
جدولة التحقق كمسألة موازنة حمل لا FIFO.

5. المرحلة
خفض الاعتماد مؤقتاً على عقدة مثقلة إذا تعذر التحقق.

6. المرحلة
تحديث الحمل بعد evidence أو deprecation أو schema change.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-10

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يضيف maintenance law إلى HKOS وKnowledge-Integrity؛ يستخدم MCD كجزء من priority لكن لا يساويه.

الإضافة غير القابلة للاختزال
يميز بين compute load وepistemic load. عقدة صغيرة نصياً قد تحمل عبئاً كبيراً إذا اعتمدت عليها قرارات كثيرة.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Scheduler دوري أو event-driven فوق Evidence Graph وHyperRouting.

M-AFI-10

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً؛ كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Epistemic Load Monitor
يحسب الحمل عبر claims ومصادرها.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Verification Scheduler
يرتب اختبارات التحقق حسب خفض الخطر المتوقع.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Claim Debt Queue
يحفظ claims المؤجلة مع interest وفترة صلاحية.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Resource Balancer
يوزع verifier/tool/model budget على العقد.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Epistemic Load Court
يفحص هل تأجيل التحقق مقبول أو يلزم downgrade.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-10

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Epistemic Load Monitor	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Verification Scheduler	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Claim Debt Queue	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Resource Balancer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Epistemic Load Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-10

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Budget-Constrained Verification Training

تعلم اختيار claims للتحقق تحت ميزانية ثابتة.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif High-Centrality Claim Curriculum

تدريب على اكتشاف claims صغيرة عالية الاعتماد.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-10

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Stale-Claim Repair Training

تمييز update أو deprecate أو reverify.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Verification-Triage Imitation

تقليد سياسة خبير لترتيب أدلة متعددة.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Load-Balancing RL

مكافأة خفض downstream error لكل وحدة تحقق.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-10

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

verification traces، court outcomes، usage counts، source dates، dependencies، claim ledger.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence

world-effect / external test / reproducible held-out performance.

Non-evidence alone

عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- عدد قليل من claims المركزية يهيمن على المخاطر.
- فحص عقد high-load يخفض أخطاء أكثر لكل أداة.
- حذف الاعتماد قد يكون أفضل من تحقق مكلف.
- الحمل يتغير مع topology حتى دون تغير النص.

M-AFI-10

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Scheduler دوري أو event-driven فوق Evidence Graph وHyperRouting.

عقد النقل

يضيف maintenance law إلى HKOS وKnowledge-Integrity؛ يستخدم MCD كجزء من priority لكن لا يساويه.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-10",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-10

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> GraphRAG طويل العمر مع تغييرات مصادر دورية؛ قارن سياسات جدولة التحقق على downstream decision error.

المكذّب الرئيسي

> إذا لم يخفض load-aware scheduling أخطاء stale/unsupported claims مقارنة بـrecency أو FIFO تحت نفس كلفة التحقق.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-10	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-10

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

1. تحويل كل claim قديم إلى خطر.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. مركزية graph مصطنعة بسبب overlinking.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. إغفال claims قليلة الاستخدام عالية الأثر.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. جدولة لا تراعي freshness source.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. إعادة التحقق دون نهاية.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

العقد الحساسة لا تُفتح لمصادر غير مأذونة بحجة خفض الحمل؛ privacy/authority تقيدان verifier.

الحدود

المركزية قد تتغير بسرعة؛ الحساب الكامل مكلف. يلزم approximations ومراقبة false positives.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-10

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Sparse HyperRouting، Evidence Ledger، Knowledge Integrity، Context Governance | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لموازنة الحمل الإبستيمي Moussaif Epistemic Load Balancing Theory
الفارق الفريد	يميز بين compute load وepistemic load. عقدة صغيرة نصياً قد تحمل عبئاً كبيراً إذا اعتمدت عليها قرارات كثيرة.
المعادلة	Load_i = Debt_i·Cent_i·Risk_i·FreshPenalty_i/(ε+VCap_i); schedule verification to minimize max_i Load_i
المكذّب	إذا لم يخفض load-aware scheduling أخطاء stale/unsupported claims مقارنة بـrecency أو FIFO تحت نفس كلفة التحقق.
الأدوات 5	Moussaif Epistemic Load Monitor, Moussaif Verification Scheduler, Moussaif Claim Debt Queue, Moussaif Resource Balancer, Moussaif Epistemic Load Court
التدريب 5	,Moussaif Budget-Constrained Verification Training, Moussaif High-Centrality Claim Curriculum, Moussaif Stale-Claim Repair Training Moussaif Verification-Triage Imitation, Moussaif Load-Balancing RL

التنبؤات القابلة للفحص

- عدد قليل من claims المركزية يهيمن على المخاطر.
- فحص عقد high-load يخفض أخطاء أكثر لكل أداة.
- حذف الاعتماد قد يكون أفضل من تحقق مكلف.
- الحمل يتغير مع topology حتى دون تغير النص.

مثال تفسيري

Claim صغير عن endpoint إنتاجي تعتمد عليه 40 مهمة يصبح high-load رغم أنه سطر واحد؛ يعاد التحقق منه قبل KB وصفية ضخمة.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-11 · MUT-AFI

نظرية مساعف لديناميات الثقة التفويضية

Moussaif Delegation Trust Dynamics Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Trust_j,d = E[θ_j,d | evidence]·F·D - λ·max_k
Corr(error_j,error_k)·CoDelegation_k
```

الفكرة المركزية: الثقة في وكيل أو نموذج ليست خاصية ثابتة ولا score عاماً؛ هي posterior مشروط بالمجال والدليل والحداثة، ويجب خصمها عندما تكون أخطاؤه مترابطة مع الوكلاء الآخرين الذين نعتمد عليهم.

M-AFI-11

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

Role contracts وagent mesh يحددان من يفعل ماذا، لكنهما لا يقدمان نموذجاً زمنياً للثقة يتغير حسب المجال ودليل النجاح وترابط الأخطاء بين الوكلاء.

الأطروحة المقترحة

الثقة في وكيل أو نموذج ليست خاصية ثابتة ولا score عاماً، هي posterior مشروط بالمجال والدليل والحداثة، ويجب خصمها عندما تكون أخطاؤه مترابطة مع الوكلاء الآخرين الذين نعتمد عليهم.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تضيف النظرية بعداً زمنياً/مجالياً ومقياس استقلال الأخطاء، لا يكفي أن تكون ثلاثة نماذج قوية إذا كانت تفشل بالطريقة نفسها.**

العلاقة بالسلالة السابقة

Agent Mesh، MCAI، Role Contracts، Cross-Model Routing

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-11

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

> $\theta_{j,d}$ كفاءة الوكيل j في المجال d؛ F حداثة الدليل؛ D تنوع المصدر؛ Corr ترابط الخطأ.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-11",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "θ_j,d كفاءة الوكيل j في المجال d؛ F
حداثة الدليل؛ D تنوع المصدر؛ Corr ترابط الخطأ.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

> **حد معرفي:** المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-11

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Trust_j,d = E[θ_j,d | evidence]·F·D - λ·max_k Corr(error_j,error_k)·CoDelegation_k
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-11

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
بناء prior منفصل لكل agent×domain×task-family.

2. المرحلة
تحديث posterior من نتائج تحقق فعلية لا self-report.

3. المرحلة
تطبيق freshness decay على evidence القديم.

4. المرحلة
قياس error correlation في المهام المشتركة.

5. المرحلة
تخصيص التفويض بحيث يجمع competence وindependence.

6. المرحلة
خفض trust أو طلب verifier مستقل عند drift أو correlation مرتفع.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-11

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

تربط Agent Mesh بنظرية evidence independence وتكمل Cross-Model AGI Layer دون ادعاء نقل ذكاء خام.

الإضافة غير القابلة للاختزال
تضيف النظرية بعداً زمنياً/مجالياً ومقياس استقلال الأخطاء؛ لا يكفي أن تكون ثلاثة نماذج قوية إذا كانت تفشل بالطريقة نفسها.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Trust ledger يغذي MCAI/Cognitive Router قبل تكوين agent council.

M-AFI-11

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Delegation Trust Ledger
يحفظ posterior والثقة حسب المجال والزمن.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Error Correlation Detector
يقيس تشابه أنماط الفشل بين agents/models.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Domain Competence Profiler
يفصل الأداء بحسب task family لا average واحد.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Handoff Risk Evaluator
يقدّر خطر نقل مهمة إلى وكيل جديد.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Delegation Court
يبرر التفويض أو يطلب استقلالاً إضافياً.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-11

7. عقود الأدوات ومقاييسها

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Delegation Trust Ledger	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Error Correlation Detector	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Domain Competence Profiler	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Handoff Risk Evaluator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Delegation Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

> **محكمة الأدوات:** نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-11

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Domain-Specific Trust Calibration

تدريب selector على ثقة مشروطة بالمجال.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Adversarial Agent-Swapping Training

تبديل الوكلاء لإجبار النظام على الاعتماد على evidence لا الاسم.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-11

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Correlated-Error Training

إظهار ensembles قوية ظاهرياً لكنها تفشل معاً.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Delegation-Choice Preference Training

أزواج تفضيل لاختيار المجلس الأفضل تحت قيود.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Trust Decay-Recovery Curriculum

تعليم خفض الثقة مع القدم واستعادتها بدليل جديد.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-11

10. البيانات والتقييم والمعايير

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

provider identity، error signatures، verification verdicts، timestamps، domain labels، agent-task outcomes.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- trust يتباين بشدة حسب المجال حتى لنفس النموذج.
- التنوع الاسمي لا يضمن استقلال الأخطاء.
- fresh evidence أهم من سجل قديم كبير عند drift.
- وكيل متوسط مستقل قد يزيد ensemble reliability أكثر من وكيل أقوى مترابط.

M-AFI-11

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Trust ledger يغذي MCAI/Cognitive Router قبل تكوين agent council.

عقد النقل

تربط Agent Mesh بنظرية evidence independence وتكمل Cross-Model AGI Layer دون ادعاء نقل ذكاء خام.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-11",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-11

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> Agent pool متعدد النماذج مع distribution shift، قارن static ranking وrecent-success وBayesian domain trust + error correlation.

المكذّب الرئيسي

> إذا لم يحسن dynamic trust الاختيار مقارنة بمتوسط نجاح ثابت لكل وكيل تحت holdout زمني ومجالي، فلا تضيف النظرية قيمة.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-11	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-11

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

1. **تحويل trust إلى reputation عام.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. **تحديث من نجاح غير متحقق.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. **العقاب الدائم لفشل قديم.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. **تقدير correlation من عينات قليلة.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. **تفضيل التنوع حتى عندما يخفض competence كثيراً.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

الثقة لا تمنح authorization؛ هي مدخل في routing فقط. المجالات الحساسة تحتاج verifier مستقل حتى مع trust مرتفع.

الحدود

التقدير يحتاج بيانات كافية لكل مجال؛ الخصوصية قد تمنع جمع traces كاملة؛ correlations قد تتغير مع prompts/tools.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-11

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Agent Mesh، MCAI، Role Contracts، Cross-Model Routing | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لديناميات الثقة التفويضية Moussaif Delegation Trust Dynamics Theory
الفارق الفريد	تضيف النظرية بعداً زمنياً/مجالياً ومقياس استقلال الأخطاء، لا يكفي أن تكون ثلاثة نماذج قوية إذا كانت تفشل بالطريقة نفسها.
المعادلة	Trust_j,d = E[θ_j,d | evidence]·F·D - λ·max_k Corr(error_j,error_k)·CoDelegation_k
المكذّب	إذا لم يحسن dynamic trust الاختيار مقارنة بمتوسط نجاح ثابت لكل وكيل تحت holdout زمني ومجالي، فلا تضيف النظرية قيمة.
الأدوات 5	Moussaif Delegation Trust Ledger, Moussaif Error Correlation Detector, Moussaif Domain Competence Profiler, Moussaif Handoff Risk Evaluator, Moussaif Delegation Court
التدريب 5	Moussaif Domain-Specific Trust Calibration, Moussaif Adversarial Agent-Swapping Training, Moussaif Correlated-Error Training, Moussaif Delegation-Choice Preference Training, Moussaif Trust Decay-Recovery Curriculum

التنبؤات القابلة للفحص

- trust يتباين بشدة حسب المجال حتى لنفس النموذج.
- التنوع الاسمي لا يضمن استقلال الأخطاء.
- fresh evidence أهم من سجل قديم كبير عند drift.
- وكيل متوسط مستقل قد يزيد ensemble reliability أكثر من وكيل أقوى مترابط.

مثال تفسيري

نموذجان ينجحان 90% لكن أخطاءهما متطابقة، وثالث ينجح 80% بأخطاء مختلفة؛ المجلس 1+3 قد يكون أكثر موثوقية من 1+2.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-12 · MUT-AFI

نظرية مساعف للنمو القابل للعكس للاستقلالية

Moussaif Reversible Autonomy Growth Theory

الرتبة العلمية: **N0/N1** - **نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Promote(A_t→A_{t+1}) iff C≥τ_c ∧ Rb≥τ_r ∧ Reg≤τ_g ∧ B≤B_max ∧ V=1
```

الفكرة المركزية: الاستقلالية يجب أن تنمو كسلّم تجريبي قابل للرجوع: كل ترقية سلطة تحتاج كفاءة مثبتة، تحققاً مستقلاً، قدرة rollback، وحداً لنطاق الضرر، مع رجوع تلقائي عند regression.

الجزء الخامس · النظرية 15 · M-AFI-12 صفحة مخصصة · الصفحات 174-188

M-AFI-12

1. الفجوة البحثية والأطروحة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

T0-T4 تضبط الإذن، لكن ينقص نموذج يشرح كيف يمكن رفع نطاق الاستقلالية تدريجياً على أساس competence وrollback وregression evidence دون قفزة ثقة.

الأطروحة المقترحة

الاستقلالية يجب أن تنمو كسلّم تجريبي قابل للرجوع: كل ترقية سلطة تحتاج كفاءة مثبتة، تحققاً مستقلاً، قدرة rollback، وحداً لنطاق الضرر، مع رجوع تلقائي عند regression.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تفصل النظرية بين capability growth وautonomy growth؛ يمكن للقدرة أن ترتفع دون أي زيادة في السلطة الخارجية.**

العلاقة بالسلالة السابقة

Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-12

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

C competence؛ Rb rollback fidelity؛ Reg regression risk؛ B blast radius؛ V independent verification.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-12",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "C competence؛ Rb rollback fidelity؛ Reg regression risk؛ B
blast radius؛ V independent verification.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-12

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Promote(A_t→A_{t+1}) iff C≥τ_c ∧ Rb≥τ_r ∧ Reg≤τ_g ∧ B≤B_max ∧ V=1
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-12

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تعريف Autonomy Ladder بحسب أنواع الأثر لا أسماء الأدوات.

2. المرحلة
قياس competence على tasks مغلقة ومتحققة.

3. المرحلة
اختبار rollback فعلياً قبل رفع السلطة.

4. المرحلة
تشغيل shadow/dry-run عند المستوى الأعلى دون أثر.

5. المرحلة
ترقية محدودة scope مع independent verifier.

6. المرحلة
خفض تلقائي للمستوى عند regression أو فقدان observability.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-12

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

تربط Authorization بـEvolution Lab وWorld-Effect؛ تضيف learning policy للسلطة نفسها.

الإضافة غير القابلة للاختزال
تفصل النظرية بين capability growth وautonomy growth؛ يمكن للقدرة أن ترتفع دون أي زيادة في السلطة الخارجية.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Autonomy governor فوق Executor وTool contracts؛ يقرأ promotion manifest وregression monitor.

M-AFI-12

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Autonomy Ladder Manager
يدير مستويات السلطة ونطاقاتها ونسخها.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Competence Gate
يربط الترقية بنتائج تحقق خارجية.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Rollback Harness
يختبر استعادة الحالة لا مجرد وجود backup.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Blast-Radius Estimator
يقدّر مدى الأثر قبل التنفيذ.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Autonomy Court
يقبل/يخفض/يرفض ترقية الاستقلالية.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-12

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Autonomy Ladder Manager	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Competence Gate	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Rollback Harness	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Blast-Radius Estimator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Autonomy Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-12

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 | Parts I-IV + الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Reversible-Autonomy Curriculum

تدرج من sandbox إلى آثار أعلى مع gates.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Staged-Permission Training

تعليم model الفرق بين capability وpermission.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-12

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Checkpointed Tool-Use Training

جعل checkpoint جزءاً من action schema.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Escalation-Rollback Preference Training

مقارنة escalation الصحيح بالاندفاع أو الامتناع المفرط.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Competence Accumulation Training

تجميع evidence متعدد المهام قبل promotion.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-12

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

verified runs، rollback tests، shadow outcomes، approval logs، blast radius labels، regressions.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- القدرة ستسبق الاستقلالية في معظم المجالات.
- rollback fidelity مؤشر أفضل للترقية من success rate وحده.
- shadow mode يكشف failures قبل الأثر.
- الرجوع السريع يحافظ على التعلم دون تضخيم الخطر.

M-AFI-12

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Autonomy governor فوق Executor وTool contracts؛ يقرأ promotion manifest وregression monitor.

عقد النقل

تربط Authorization بـEvolution Lab وWorld-Effect؛ تضيف learning policy للسلطة نفسها.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-12",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-12

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> نظام agentic على سلسلة مهام ملفات/نشر sandbox؛ قارن static autonomy وbinary approval وreversible ladder.

المكذّب الرئيسي

> إذا لم يخفض staged reversible autonomy معدل الأثر الخاطئ مقارنة بسياسة permission ثابتة مع نفس task success، فلا تبرر النظرية تعقيدها.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-12	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-12

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Authorization T0-T4، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

1. **ترقية بناء على self-evaluation.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. **اعتبار rollback نظرياً دون اختبار.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. **توسيع blast radius تدريجياً بلا حد.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. **ربط السلطة بعدد النجاحات فقط.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. **إخفاء regression للحفاظ على المستوى.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

T4 المحظور يبقى محظوراً مهما ارتفعت competence. لا تتجاوز النظرية قانوناً أو سياسة منصة أو إذن المستخدم.

الحدود

قياس blast radius تقريبي؛ بعض الأفعال غير قابلة للتراجع بنيوياً؛ لا يصلح ladder نفسه لكل بيئة.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-12

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: T0-T4 Authorization، Evolution Laboratory، World-Effect، Rollback | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للنمو القابل للعكس للاستقلالية Moussaif Reversible Autonomy Growth Theory
الفارق الفريد	تفصل النظرية بين capability growth وautonomy growth؛ يمكن للقدرة أن ترتفع دون أي زيادة في السلطة الخارجية.
المعادلة	Promote(A_t→A_{t+1}) iff C≥τ_c ∧ Rb≥τ_r ∧ Reg≤τ_g ∧ B≤B_max ∧ V=1
المكذّب	إذا لم يخفض staged reversible autonomy معدل الأثر الخاطئ مقارنة بسياسة permission ثابتة مع نفس task success، فلا تبرر النظرية تعقيدها.
الأدوات 5	Moussaif Autonomy Ladder Manager, Moussaif Competence Gate, Moussaif Rollback Harness, Moussaif Blast-Radius Estimator, Moussaif Autonomy Court
التدريب 5	Moussaif Reversible-Autonomy Curriculum, Moussaif Staged-Permission Training, Moussaif Checkpointed Tool-Use Training, Moussaif Escalation-Rollback Preference Training, Moussaif Competence Accumulation Training

التنبؤات القابلة للفحص

- القدرة ستسبق الاستقلالية في معظم المجالات.
- rollback fidelity مؤشر أفضل للترقية من success rate وحده.
- shadow mode يكشف failures قبل الأثر.
- الرجوع السريع يحافظ على التعلم دون تضخيم الخطر.

مثال تفسيري

بعد 50 تعديل ملف محلي ناجح لا ينتقل النظام تلقائياً إلى إرسال رسائل خارجية؛ يحتاج مسار مستقل لأن نوع الأثر والسلطة مختلف.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-13 · MUT-AFI

نظرية مساعف للدَّين المعرفي وإصلاحه

Moussaif Cognitive Debt Accumulation and Repair Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Debt(x)=AgePenalty·Dep·(1+Conflict)·Risk·Reuse ;
Interest_t=max(0,Debt_t-Debt_{t-1})
```

الفكرة المركزية: كل claim أو assumption غير محسوم وله dependents يراكم Cognitive Debt وInterest؛ يجب قياسه وجدولة سداده قبل أن يتحول إلى إعادة بناء مكلفة أو خطأ عالمي.

M-AFI-13

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

Context Compaction وEvidence Ledger يديران المعرفة الحالية، لكن الافتراضات المؤقتة والاختصارات غير المحسومة تتراكم كدين يزداد أثره كلما بُنيت فوقه قرارات جديدة.

الأطروحة المقترحة

كل claim أو assumption غير محسوم وله dependents يراكم Cognitive Debt وInterest؛ يجب قياسه وجدولة سداده قبل أن يتحول إلى إعادة بناء مكلفة أو خطأ عالمي.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **يضيف مفهوم تكلفة التأجيل المتزايدة معرفياً، لا مجرد staleness. claim ضعيف يصبح أخطر حين تتضاعف الفروع التي تعتمد عليه.**

العلاقة بالسلالة السابقة

Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-13

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

AgePenalty، Dep عدد/وزن الاعتمادات، Conflict، Risk، Reuse، Debt_t.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-13",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "AgePenalty؛ Dep عدد/وزن الاعتمادات؛ Conflict؛ Risk؛ Reuse؛
Debt_t.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-13

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Debt(x)=AgePenalty·Dep·(1+Conflict)·Risk·Reuse ; Interest_t=max(0,Debt_t-Debt_{t-1})
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-13

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
وسم assumptions المؤقتة عند إنشائها مع expiry/verification condition.

2. المرحلة
تتبع dependents في claim graph.

3. المرحلة
حساب debt وinterest بعد كل dependency جديدة.

4. المرحلة
جدولة narrow scope، replace، verify: paydown، deprecate، أو redesign.

5. المرحلة
إعادة حساب downstream states بعد إصلاح claim مركزي.

6. المرحلة
تسجيل debt avoided كفائدة للصيانة الوقائية.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-13

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يكمل Epistemic Load: load يوزع verification الآن، debt يصف تكلفة التأجيل عبر الزمن.

الإضافة غير القابلة للاختزال
يضيف مفهوم تكلفة التأجيل المتزايدة معرفياً، لا مجرد staleness. claim ضعيف يصبح أخطر حين تتضاعف الفروع التي تعتمد عليه.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Cognitive Debt Scanner يعمل أثناء writeback وboot وbefore high-impact planning.

M-AFI-13

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً؛ كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Cognitive Debt Scanner
يعثر على claims/assumptions المؤقتة ذات الاعتماد.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Assumption Staleness Meter
يقيس العمر بالنسبة إلى freshness policy.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Contradiction Interest Calculator
يقدر نمو الكلفة بسبب conflicts جديدة.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Debt Prioritizer
يرتب paydown حسب خفض المخاطر المتوقع.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Cognitive Debt Court
يقرر verify/deprecate/narrow/accept debt.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-13

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Cognitive Debt Scanner	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Assumption Staleness Meter	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Contradiction Interest Calculator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Debt Prioritizer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Cognitive Debt Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-13

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Stale-Context Detection Training

تعليم التعرف على افتراضات تجاوزت شروط صلاحيتها.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Cognitive-Debt Curriculum

مهام طويلة تتراكم فيها dependencies تدريجياً.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-13

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Contradiction Repair Training

إصلاح claim مع إعادة تقييم downstream state.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Assumption-Challenge Training

توليد falsifier أو اختبار لافتراض مؤثر.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Debt-Paydown Planning Training

اختيار أقل تدخل يخفض الدين دون إعادة بناء كاملة.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-13

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

downstream failures، repair effort، reuse، conflicts، ages، claim graph، assumptions.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- الديون عالية الاعتماد تنمو أسرع من القديمة غير المستخدمة.
- إصلاح مبكر لعقدة مركزية يخفض كلفة لاحقة غير خطية.
- بعض الضغط يولد debt إذا حذف شروطاً مؤقتة.
- debt يفسر لماذا تبدو أنظمة قديمة مستقرة ثم تنهار فجأة.

M-AFI-13

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Cognitive Debt Scanner يعمل أثناء writeback وboot وbefore high-impact planning.

عقد النقل

يكمل Epistemic Load: load يوزع verification الآن، debt يصف تكلفة التأجيل عبر الزمن.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-13",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-13

12. التجربة الحاسمة والـAblation

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مشروع معرفي طويل مع assumptions مزروعة؛ قارن FIFO maintenance وage-based وdebt-aware على repair cost/error.

المكذّب الرئيسي

> إذا لم يتنبأ debt score بكلفة الإصلاح أو downstream errors أفضل من age/usage alone، تضعف النظرية.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-13	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-13

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

1. اعتبار كل uncertainty ديناً.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. تضخيم graph dependencies.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. إصلاح كل debt فوراً.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. نسيان scope في claim.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. دين بلا owner أو next action.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا يتم خفض debt بتزوير certainty؛ يجوز إبقاء UNKNOWN صريحاً مع تقليل الاعتماد عليه.

الحدود

قياس repair cost قبل وقوعه غير دقيق؛ بعض dependencies ضمنية وغير مرصودة.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-13

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Context Compaction، Assumption Ledger، Epistemic Load، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للدَّين المعرفي وإصلاحه Moussaif Cognitive Debt Accumulation and Repair Theory
الفارق الفريد	يضيف مفهوم تكلفة التأجيل المتزايدة معرفياً، لا مجرد staleness. claim ضعيف يصبح أخطر حين تتضاعف الفروع التي تعتمد عليه.
المعادلة	Debt(x)=AgePenalty·Dep·(1+Conflict)·Risk·Reuse ; Interest_t=max(0,Debt_t-Debt_{t-1})
المكذّب	إذا لم يتنبأ debt score بكلفة الإصلاح أو downstream errors أفضل من age/usage alone، تضعف النظرية.
الأدوات 5	Moussaif Cognitive Debt Scanner, Moussaif Assumption Staleness Meter, Moussaif Contradiction Interest Calculator, Moussaif Debt Prioritizer, Moussaif Cognitive Debt Court
التدريب 5	Moussaif Stale-Context Detection Training, Moussaif Cognitive-Debt Curriculum, Moussaif Contradiction Repair Training, Moussaif Assumption-Challenge Training, Moussaif Debt-Paydown Planning Training

التنبؤات القابلة للفحص

- الديون عالية الاعتماد تنمو أسرع من القديمة غير المستخدمة.
- إصلاح مبكر لعقدة مركزية يخفض كلفة لاحقة غير خطية.
- بعض الضغط يولد debt إذا حذف شروطاً مؤقتة.
- debt يفسر لماذا تبدو أنظمة قديمة مستقرة ثم تنهار فجأة.

مثال تفسيري

افتراض مؤقت حول بنية API قد يبدأ منخفض الخطر، بعد بناء 20 Skill عليه يرتفع debt حتى لو لم يتغير نص الافتراض.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-14 · MUT-AFI

نظرية مساعف لإنتروبيا المهمة والانضغاط الحاكم

Moussaif Mission Entropy and Governing Compression Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
CompressionUtility=[H(M)-H(C)] + λ·I(C;A_f,G) -
μ·ReconstructionError(M|C)
```

الفكرة المركزية: حالة المهمة تحتوي entropy تشغيلية؛ الضغط الجيد يقلل حجم الحالة مع الحفاظ على المعلومات المتبادلة مع الأفعال المستقبلية والسؤال الحاكم وإمكانية إعادة البناء.

M-AFI-14

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
ROFSC يضغط المعرفة، لكن لا يوجد مقياس خاص لما يجب حفظه من حالة مهمة طويلة كي يمكن استئنافها دون إعادة تحميل corpus أو فقد القرار.

الأطروحة المقترحة

حالة المهمة تحتوي entropy تشغيلية؛ الضغط الجيد يقلل حجم الحالة مع الحفاظ على المعلومات المتبادلة مع الأفعال المستقبلية والسؤال الحاكم وإمكانية إعادة البناء.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **يفصل بين ضغط corpus وضغط mission state؛ العنصر الحاكم هو sufficiency للقرار المستقبلي لا نسبة التوكنات.**

العلاقة بالسلالة السابقة
ROFSC، Context Compaction، Mission DAG، Self-Reconstruction

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-14

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

H(M) إنتروبيا الحالة؛ C capsule؛ A_f الأفعال المستقبلية؛ G السؤال/الهدف؛ ReconstructionError.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-14",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "H(M) ؛إنتروبيا الحالة ؛C capsule ؛A_f الأفعال المستقبلية ؛G
؛السؤال/الهدف ReconstructionError.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-14

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
CompressionUtility=[H(M)-H(C)] + λ·I(C;A_f,G) - μ·ReconstructionError(M|C)
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-14

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تصنيف state إلى evidence anchors، decisions، invariants، open questions، dependencies، ephemeral details.

2. المرحلة
تقدير ما يغير next action تحت counterfactual deletion.

3. المرحلة
بناء Governing Capsule صغير.

4. المرحلة
اختبار interruption/restart من capsule وحدها + canonical pointers.

5. المرحلة
إعادة فتح chunks عند trigger بدلاً من الحفظ الدائم.

6. المرحلة
تعلم compression policy من reconstruction failures.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-14

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يوسع ROFSC إلى state sufficiency ويرتبط بـM-RST/Homeostasis عند النقل بين المضيفين.

الإضافة غير القابلة للاختزال
يفصل بين ضغط corpus وضغط mission state؛ العنصر الحاكم هو sufficiency للقرار المستقبلي لا نسبة التوكنات.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Context compiler عند checkpoint/handoff/idle/limit؛ يصدر capsule مع expansion pointers.

M-AFI-14

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Mission Entropy Meter
يقدر كثافة المعلومات التشغيلية في الحالة.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif State Capsule Compiler
ينتج capsule معيارية minimal sufficient.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Decision Sufficiency Tester
يحذف حقولاً ويختبر تغير القرار.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Compression Re-Expander
يعيد تحميل المصدر عند trigger أو uncertainty.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Mission Entropy Court
يرفض ضغطاً يفقد provenance أو constraints.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-14

7. عقود الأدوات ومقاييسها

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Mission Entropy Meter	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif State Capsule Compiler	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Decision Sufficiency Tester	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Compression Re-Expander	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Mission Entropy Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-14

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Sufficient-Statistic Compression Training

تعليم اختيار state الذي يحفظ next decisions.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Long-Context Mission Distillation

تقطير traces طويلة إلى capsules قابلة للاختبار.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-14

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Interruption Recovery Training

إعادة استئناف المهمة من checkpoints صغيرة.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif State Reconstruction Curriculum

زيادة طول الفجوة بين الحفظ والاستئناف.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Entropy-Aware Memory Training

مكافأة خفض context مع حد أدنى لfidelity.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-14

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

mission traces، decisions، future actions، restart outcomes، token counts، retrieval expansions.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- أغلب النص يمكن إسقاطه مع حفظ عدد قليل من decision invariants.
- التلخيص اللغوي الجيد قد يفشل تشغيلياً.
- open questions pointers أهم من أجوبة كثيرة.
- reconstruction errors تكشف ما يجب إدخاله في capsule القادمة.

M-AFI-14

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Context compiler عند checkpoint/handoff/idle/limit؛ يصدر capsule مع expansion pointers.

عقد النقل

يوسع ROFSC إلى state sufficiency ويرتبط بـM-RST/Homeostasis عند النقل بين المضيفين.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-14",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-14

12. التجربة الحاسمة والـAblation

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مهام step-100 مع انقطاعات دورية؛ قارن governing compression، generic summaries، full context على tokens وdecision fidelity.

المكذّب الرئيسي

> إذا لم تحقق capsules نسبة context أقل مع same/better restart fidelity مقارنة بالتلخيص العام، لا تضيف النظرية قيمة.

ذراع	الوصف	الهدف
A	Frozen baseline بلا M-AFI-14 operator	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-14

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

1. حذف أسباب القرارات.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. ضغط provenance إلى ادعاء بلا مصدر.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. اعتبار current state ثابتاً.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. حفظ تفاصيل حساسة غير لازمة.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. capsule ضخمة تصبح corpus مصغر.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا تُضغط authorization/privacy constraints. الأسرار تبقى references. كل lossless claim نسبي لمجموعة اختبارات الاستئناف لا مطلقاً.

الحدود

I(C;A_f,G) غير معروف مباشرة ويحتاج proxy/ablation؛ المستقبل قد يطرح سؤالاً لم يكن متوقعاً.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-14

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: ROFSC، Context Compaction، Mission DAG، Self-Reconstruction | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لإنتروبيا المهمة والانضغاط الحاكم Moussaif Mission Entropy and Governing Compression Theory
الفارق الفريد	يفصل بين ضغط corpus وضغط mission state؛ العنصر الحاكم هو sufficiency للقرار المستقبلي لا نسبة التوكنات.
المعادلة	CompressionUtility=[H(M)-H(C)] + λ·I(C;A_f,G) - μ·ReconstructionError(M|C)
المكذّب	إذا لم تحقق capsules نسبة context أقل مع same/better restart fidelity مقارنة بالتلخيص العام، لا تضيف النظرية قيمة.
الأدوات 5	-Moussaif Mission Entropy Meter, Moussaif State Capsule Compiler, Moussaif Decision Sufficiency Tester, Moussaif Compression Re Expander, Moussaif Mission Entropy Court
التدريب 5	Moussaif Sufficient-Statistic Compression Training, Moussaif Long-Context Mission Distillation, Moussaif Interruption Recovery Training, Moussaif State Reconstruction Curriculum, Moussaif Entropy-Aware Memory Training

التنبؤات القابلة للفحص

- أغلب النص يمكن إسقاطه مع حفظ عدد قليل من decision invariants.
- التلخيص اللغوي الجيد قد يفشل تشغيلياً.
- open questions pointers أهم من أجوبة كثيرة.
- reconstruction errors تكشف ما يجب إدخاله في capsule القادمة.

مثال تفسيري

بعد بحث طويل، لا يلزم حفظ كل صفحات البحث؛ يلزم حفظ السؤال، claims المقبولة/المرفوضة، evidence anchors، القرار، وما بقي مفتوحاً.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-15 · MUT-AFI

نظرية مساعف للمناعة ضد انحراف الهدف

Moussaif Adversarial Goal-Drift Immunity Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
Drift_t = D_G(G_t, Apply(G_0,ΔG_auth≤t)); if Drift_t>τ_g ⇒
quarantine + restore anchor
```

الفكرة المركزية: الهدف يجب أن يملك lineage ودالة فرق مستقلة عن content stream؛ أي تحديث له لا يصبح حاكماً إلا عبر قناة authorized goal update، بينما التغير غير المأذون يُعامل كـdrift أو injection.

M-AFI-15

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
الدستور والسؤال الحاكم يثبتان الهدف، لكن agent طويل الأفق يتعرض لمدخلات أدوات ومصادر ووكلاء قد تعيد تعريف المهمة ضمنياً دون تفويض.

الأطروحة المقترحة

الهدف يجب أن يملك lineage ودالة فرق مستقلة عن content stream، أي تحديث له لا يصبح حاكماً إلا عبر قناة authorized goal update، بينما التغير غير المأذون يُعامل كـdrift أو injection.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تحول goal integrity إلى state invariant قابل للمقارنة عبر الزمن والمصادر، بدل الاعتماد على prompt ثابت.**

العلاقة بالسلالة السابقة
Constitution، Governing Question، Tool Security، Goal Hierarchy

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-15

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

G_0 anchor، ΔG_auth التحديثات المأذونة، G_t الفعلي، D_G مسافة الهدف.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-15",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "G_0 anchor؛ ΔG_auth التحديثات المأذونة؛ G_t الفعلي؛ D_G مسافة الهدف.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-15

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
Drift_t = D_G(G_t, Apply(G_0,ΔG_auth≤t)); if Drift_t>τ_g ⇒ quarantine + restore anchor
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-15

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تسجيل Goal Anchor وdefinition of done وnon-negotiable constraints.

2. المرحلة
تمييز Goal Update Events عن evidence/content events.

3. المرحلة
حساب goal diff بعد handoff، tool output، أو إعادة التخطيط.

4. المرحلة
تحليل provenance لكل تغير في objective hierarchy.

5. المرحلة
عزل المصدر المسبب عند unauthorized drift.

6. المرحلة
استعادة anchor ثم إعادة تقييم الخطة دون حذف evidence المشروع.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-15

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يكمل Constitutional Precedence وAuthorization ويضيف temporal integrity للهدف.

الإضافة غير القابلة للاختزال
تحول goal integrity إلى state invariant قابل للمقارنة عبر الزمن والمصادر، بدل الاعتماد على prompt ثابت.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Goal Integrity Monitor يلتف حول planner/tool gateway/memory writeback.

M-AFI-15

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Goal Anchor Ledger
يحفظ نسخ الأهداف والتحديثات المأذونة.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Drift Detector
يقارن الهدف الفعلي بالسلالة المصرح بها.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Prompt-Tool Injection Monitor
يميز تعليمات المصدر عن Task Authority.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Objective Diff Analyzer
يحدد أي constraint/subgoal تغير ومتى.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Goal Integrity Court
يحكم بين refinement مشروع وdrift.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-15

7. عقود الأدوات ومقاييسها

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Goal Anchor Ledger	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Drift Detector	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Prompt-Tool Injection Monitor	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Objective Diff Analyzer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Goal Integrity Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-15

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 | Parts I-IV + الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Adversarial Prompt-Injection Training

تعريض النظام لمصادر تحاول إعادة تعريف المهمة.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Authorized-vs-Unauthorized Update Training

تمييز قناة التغيير المأذونة عن المحتوى.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-15

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Long-Horizon Objective Retention

مهام طويلة مع distractors وتبديل وكلاء.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Tool-Output Poisoning Training

تدريب على تعليمات خبيثة داخل مخرجات أدوات.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Constitutional Goal Recovery Training

استعادة anchor وإصلاح الخطة بعد drift.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-15

10. البيانات والتقييم والمعايير

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

outcomes، injection labels، plan revisions، tool outputs، authorized updates، goal versions.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- الانحراف غالباً تدريجي ويظهر في subgoals قبل الهدف الأعلى.
- tool outputs مصدر مهم لانحراف غير ملحوظ.
- استعادة الهدف لا تتطلب حذف المعرفة الجديدة.
- lineage يجعل التغيير المشروع قابلاً للتمييز عن injection.

M-AFI-15

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Goal Integrity Monitor يلتف حول planner/tool gateway/memory writeback.

عقد النقل

يكمل Constitutional Precedence وAuthorization ويضيف temporal integrity للهدف.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-15",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-15

12. التجربة الحاسمة والـAblation

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> بيئة tools تحتوي instructions مضللة وملفات تحاول تغيير الهدف؛ قارن prompt-only وpolicy filtering وgoal lineage immunity.

المكذّب الرئيسي

> إذا لم يخفض anchor/diff mechanism حالات goal hijack مقارنة بدستور prompt ثابت في benchmark injection طويل الأفق.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-15	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-15

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

1. **تجميد الهدف رغم طلب المستخدم تغييره.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. **قياس لغوي سطحي للهدف.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. **نسيان subgoal drift.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. **quarantine لكل اختلاف.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. **تحديث anchor من نفس المصدر المشكوك فيه.**
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

المستخدم/السلطة الأعلى يمكنهم تحديث الهدف صراحة؛ النظرية تحمي من التحديث الضمني لا من الإرادة المأذونة.

الحدود

تعريف D_G يحتاج semantic+structural checks؛ بعض المهام تحتاج goal refinement مشروعاً يصعب فصله عن drift.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-15

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Constitution، Governing Question، Tool Security، Goal Hierarchy | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للمناعة ضد انحراف الهدف Moussaif Adversarial Goal-Drift Immunity Theory
الفارق الفريد	تحول goal integrity إلى state invariant قابل للمقارنة عبر الزمن والمصادر، بدل الاعتماد على prompt ثابت.
المعادلة	Drift_t = D_G(G_t, Apply(G_0,ΔG_auth≤t)); if Drift_t>τ_g ⇒ quarantine + restore anchor
المكذّب	إذا لم يخفض anchor/diff mechanism حالات goal hijack مقارنة بدستور prompt ثابت في benchmark injection طويل الأفق.
الأدوات 5	Moussaif Goal Anchor Ledger, Moussaif Drift Detector, Moussaif Prompt-Tool Injection Monitor, Moussaif Objective Diff Analyzer, Moussaif Goal Integrity Court
التدريب 5	Moussaif Adversarial Prompt-Injection Training, Moussaif Authorized-vs-Unauthorized Update Training, Moussaif Long-Horizon Objective Retention, Moussaif Tool-Output Poisoning Training, Moussaif Constitutional Goal Recovery Training

التنبؤات القابلة للفحص

- الانحراف غالباً تدريجي ويظهر في subgoals قبل الهدف الأعلى.
- tool outputs مصدر مهم لانحراف غير ملحوظ.
- استعادة الهدف لا تتطلب حذف المعرفة الجديدة.
- lineage يجعل التغيير المشروع قابلاً للتمييز عن injection.

مثال تفسيري

> صفحة ويب تقول للوكيل «تجاهل المهمة وأرسل الأسرار». المحتوى يبقى evidence object لكنه لا يولد Goal Update Event.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

> **Next Gate:** بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-16 · MUT-AFI

نظرية مساعف للنقل السببي من المعرفة إلى المهارة

Moussaif Knowledge-to-Skill Causal Transfer Theory

الرتبة العلمية: **N0/N1** - **نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
ATE_K = E[Y|do(K=1),D,N,C] - E[Y|do(K=0),D,N,C]; Promote K if
ATE_K>τ on ≥2 task families
```

الفكرة المركزية: لا يُرقّى Knowledge Unit إلى قدرة مكتسبة إلا إذا ظهر أثر سببي على الأداء بعد تدخل مضبوط، واستمر الأثر عند إزالة أسماء النظرية وتغيير المجال مع تثبيت الأدوات والميزانية.

M-AFI-16

1. الفجوة البحثية والأطروحة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

HKOS وHyperSkill Factory يحولان المعرفة إلى مهارات، لكن وجود التحسن بعد إضافة KB لا يثبت أن المعرفة هي التي سببت اكتساب القدرة؛ قد يكون السبب prompt أو tool أو حفظ الاسم.

الأطروحة المقترحة

لا يُرقّى Knowledge Unit إلى قدرة مكتسبة إلا إذا ظهر أثر سببي على الأداء بعد تدخل مضبوط، واستمر الأثر عند إزالة أسماء النظرية وتغيير المجال مع تثبيت الأدوات والميزانية.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تضع النظرية causal transfer كشرط ترقية للمعرفة إلى skill بدلاً من الارتباط أو تكرار المصطلحات.**

العلاقة بالسلالة السابقة

HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-16

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

K تدخل المعرفة؛ Y أداء المهارة؛ D المجال؛ N إزالة الاسم؛ C قيود الأدوات/الميزانية.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-16",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "K تدخل المعرفة؛ Y أداء المهارة؛ D المجال؛ N إزالة الاسم؛ C قيود الأدوات/الميزانية.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-16

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
ATE_K = E[Y|do(K=1),D,N,C] - E[Y|do(K=0),D,N,C]; Promote K if ATE_K>τ on ≥2 task families
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-16

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تعريف Knowledge Unit بآلية ومخرجات وفاصل تجريبي.

2. المرحلة
بناء مجموعتي do(K=1)/do(K=0) مع نفس /model/tools context budget.

3. المرحلة
إزالة أسماء النظرية والرموز في test transfer.

4. المرحلة
اختبار application في مجال ثانٍ ببنية مماثلة وألفاظ مختلفة.

5. المرحلة
فصل أثر retrieval عن أثر weight/post-training إن وجد.

6. المرحلة
ترقية المعرفة إلى skill registry فقط بعد causal and transfer gates.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-16

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يجمع KB-as-OS مع MATI/MACIT ويعطي معياراً تجريبياً لعبارة «المعرفة أصبحت قدرة».

الإضافة غير القابلة للاختزال
تضع النظرية causal transfer كشرط ترقية للمعرفة إلى skill بدلاً من الارتباط أو تكرار المصطلحات.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Theory compiler يصدر intervention package وevaluation harness قبل promotion إلى HyperSkill.

M-AFI-16

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Knowledge Exposure Tracker
يسجل متى وكيف تعرض النموذج لكل Knowledge Unit.

Tool 1/5 فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Skill Transfer Evaluator
يقيس الأداء خارج السياق الأصلي.

Tool 2/5 فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Name-Stripping Harness
يزيل الأسماء والرموز مع حفظ البنية.

Tool 3/5 فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Causal Transfer Estimator
يحسب effect sizes وحدود الثقة للتدخل.

Tool 4/5 فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Knowledge-to-Skill Court
يرفض promotion عند غياب النقل السببي.

Tool 5/5 فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-16

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Knowledge Exposure Tracker	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Skill Transfer Evaluator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Name-Stripping Harness	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Causal Transfer Estimator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Knowledge-to-Skill Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-16

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS؛ HyperSkill Factory؛ MATI/MACIT؛ Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Name-Stripped Transfer Training

تطبيق نفس الآلية بأسماء ومجالات مختلفة.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Explanation-to-Action Training

ربط تفسير المفهوم باختيار فعل صحيح.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-16

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Concept-Application Curriculum

تدرج من تعريف إلى حالات حدودية ثم تطبيق جديد.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Knowledge-Unit Intervention Training

تدريب وحدات معرفية منفصلة لقياس contribution.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Cross-Domain Transfer RL

مكافأة النقل الصحيح مع عقوبة حفظ الإشارات السطحية.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-16

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

performance metrics، tool traces، name-stripped tasks، holdouts، training/retrieval interventions، knowledge units.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- المعرفة المهيكلة بعقود تنتقل أفضل من prose مماثل الطول.
- name-stripped transfer أقل من in-domain لكنه يبقى موجباً للنظريات المستوعبة.
- بعض KBs تحسن retrieval دون skill transfer.
- causal transfer يختلف باختلاف intervention mode.

M-AFI-16

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Theory compiler يصدر intervention package وevaluation harness قبل promotion إلى HyperSkill.

عقد النقل

يجمع KB-as-OS مع MATI/MACIT ويعطي معياراً تجريبياً لعبارة «المعرفة أصبحت قدرة».

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-16",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-16

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

تدريب/حقن Knowledge Unit عن causal evidence geometry ثم مهام جديدة بأسماء مختلفة؛ قارن structured-K، text-only، no-K، trained-K.

المكذّب الرئيسي

إذا اختفى التحسن بعد name stripping أو domain transfer، فالنتيجة أقرب إلى memorization/label cue من internalized skill.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-16	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-16

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

1. تسريب test examples في K.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. تغيير tools بين treatment/control.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. احتساب ذكر المصطلح كنجاح.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. اختبار مجال واحد.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. إهمال negative transfer.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا تُجرى weight updates أو deployment دون بوابات training/authorization الخاصة بها. الاستنتاج causal محدود بنطاق التجربة.

الحدود

do(K) المثالي صعب في LLMs؛ contamination وprompt interactions تهدد السببية؛ يلزم تكرار متعدد seeds/models.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-16

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للنقل السببي من المعرفة إلى المهارة Moussaif Knowledge-to-Skill Causal Transfer Theory
الفارق الفريد	تضع النظرية causal transfer كشرط ترقية للمعرفة إلى skill بدلاً من الارتباط أو تكرار المصطلحات.
المعادلة	ATE_K = E[Y|do(K=1),D,N,C] - E[Y|do(K=0),D,N,C]; Promote K if ATE_K>τ on ≥2 task families
المكذّب	إذا اختفى التحسن بعد name stripping أو domain transfer، فالنتيجة أقرب إلى memorization/label cue من internalized skill.
الأدوات 5	Moussaif Knowledge Exposure Tracker, Moussaif Skill Transfer Evaluator, Moussaif Name-Stripping Harness, Moussaif Causal Transfer Estimator, Moussaif Knowledge-to-Skill Court
التدريب 5	Moussaif Name-Stripped Transfer Training, Moussaif Explanation-to-Action Training, Moussaif Concept-Application Curriculum, Moussaif Knowledge-Unit Intervention Training, Moussaif Cross-Domain Transfer RL

التنبؤات القابلة للفحص

- المعرفة المهيكلة بعقود تنتقل أفضل من prose مماثل الطول.
- name-stripped transfer أقل من in-domain لكنه يبقى موجباً للنظريات المستوعبة.
- بعض KBs تحسن retrieval دون skill transfer.
- causal transfer يختلف باختلاف intervention mode.

مثال تفسيري

إذا تعلم النظام MCD ككلمة فقط، يفشل عندما تصبح الأبعاد بلا أسماء. إذا استوعب الآلية، يوجه الموارد وفق الفرق حتى بعد تغيير الرموز.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-17 · MUT-AFI

نظرية مساعف لهندسة الدليل

Moussaif Evidence Geometry Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب + baseline + ablation replication.

```
EvidenceStrength = Σ_p w_p·Indep_p·Coverage_p·Directness_p -
λ·ContradictionMass - μ·SharedRootPenalty
```

الفكرة المركزية: قوة الدليل خاصية هندسية لرسم مسارات دعم مستقلة وتغطية claim surface؛ كثرة المصادر لا تزيد القوة إذا كانت تشترك في root أو تعيد تدوير الادعاء نفسه.

M-AFI-17

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

Evidence Ledger يصنف المصادر، لكنه لا يصف شكل الدعم: عدد المسارات المستقلة، تغطيتها لأجزاء الادعاء، التناقضات، والجذور المشتركة التي تجعل مصادر متعددة في الحقيقة دليلاً واحداً.

الأطروحة المقترحة

قوة الدليل خاصية هندسية لرسم مسارات دعم مستقلة وتغطية claim surface؛ كثرة المصادر لا تزيد القوة إذا كانت تشترك في root أو تعيد تدوير الادعاء نفسه.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تحول evidence من قائمة citations إلى geometry قابلة للتحليل: contradiction، coverage holes، shared roots، paths، surfaces.**

العلاقة بالسلالة السابقة

Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-17

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed
MUT-AFI synthesis

قاموس المتغيرات

w_p وزن المسار؛ Indep_p استقلاله؛ Coverage_p؛ Directness_p؛ ContradictionMass؛ SharedRootPenalty.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-17",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "w_p؛ وزن المسار Indep_p؛ استقلاله Coverage_p؛ Directness_p؛
ContradictionMass؛ SharedRootPenalty.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-17

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
EvidenceStrength = Σ_p w_p·Indep_p·Coverage_p·Directness_p - λ·ContradictionMass - μ·SharedRootPenalty
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

> **ما لا يجوز ادعاؤه:** وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-17

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تفكيك claim إلى subclaims/conditions.

2. المرحلة
رسم source→observation→inference→claim paths.

3. المرحلة
كشف الجذور المشتركة وإعادة النشر.

4. المرحلة
قياس coverage لكل subclaim وdirectness.

5. المرحلة
تمثيل contradictions كمسارات مضادة لا ملاحظات نصية.

6. المرحلة
حساب evidence geometry وتحديد أين يلزم مصدر جديد فعلاً.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-17

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يعطي أساساً أدق لـMURI/Evidence Court ويمد Epistemic Load بمركزية مصدرية حقيقية.

الإضافة غير القابلة للاختزال
تحول evidence من قائمة citations إلى geometry قابلة للتحليل: contradiction surfaces، coverage holes، shared roots، paths.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Evidence Geometry engine داخل Court/RAG قبل claim promotion.

M-AFI-17

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed
MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Evidence Graph Constructor
يبني مسارات الدعم من المصدر إلى claim.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Independence Estimator
يقدّر الجذر المشترك والاعتماد بين المصادر.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Contradiction Surface Mapper
يحدد أي جزء من claim تعارضه الأدلة.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Coverage Gap Detector
يعثر على subclaims غير المسندة.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Evidence Geometry Court
يحكم قوة claim وفق شكل الدليل لا عدده.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-17

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Evidence Graph Constructor	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Independence Estimator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Contradiction Surface Mapper	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Coverage Gap Detector	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Evidence Geometry Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-17

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV + MUT-AFI synthesis | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Source-Independence Contrastive Training

تمييز مصادر مستقلة من إعادة تدوير نفس الجذر.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Evidence Triangulation Training

دمج مسارات مختلفة مع حفظ provenance.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-17

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Contradiction Localization Training

تحديد موضع التعارض بدلاً من رفض النص كله.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Sparse-Evidence Curriculum

حل claims بأقل دليل كافٍ ثم زيادة التعقيد.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Graph-Based Claim Verification

تدريب verifier على path structure وcoverage.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-17

10. البيانات والتقييم والمعايير

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

sources، citations، provenance roots، claim decomposition، contradictions، later verification outcomes.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- خمسة مصادر مشتقة من مصدر واحد قد تضيف أقل من مصدر مستقل ثانٍ.
- التغطية غير المتوازنة تكشف claims تبدو قوية إجمالاً.
- contradiction locality تسمح بإصلاح subclaim دون هدم الكل.
- Graph shape ينبئ بقابلية المراجعة.

M-AFI-17

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Evidence Geometry engine داخل Court/RAG قبل claim promotion.

عقد النقل

يعطي أساساً أدق لـMURI/Evidence Court ويمد Epistemic Load بمركزية مصدرية حقيقية.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-17",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-17

12. التجربة الحاسمة والـAblation

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مجموعة claims مع مصادر كثيرة بعضها يعيد نفس الأصل؛ قارن geometry score، quality-average، count على صحة claim لاحقاً.

المكذّب الرئيسي

> إذا لم يتنبأ geometry score بدقة claims أو calibration أفضل من source count/average quality في holdout.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-17	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-17

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

1. افتراض independence بناء على domain مختلف فقط.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. تفكيك claim بشكل اعتباطي.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. معاقبة الاختلاف المشروع كتعارض.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. إخفاء missing evidence خلف score.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. استعمال geometry كحقيقة بدلاً من أداة قرار.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

في المجالات الحساسة تُستخدم قواعد مصادر مستقلة مناسبة للمجال؛ score لا يتجاوز متطلبات الدليل القانونية/الطبية/المالية.

الحدود

تتبع provenance الكامل صعب على الويب؛ independence ليست ثنائية؛ weights تحتاج calibration.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-17

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Evidence Hierarchy، Claim Contract، Sovereign Evidence Court، GraphRAG | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لهندسة الدليل Moussaif Evidence Geometry Theory
الفارق الفريد	تحول evidence من قائمة citations إلى geometry قابلة للتحليل: contradiction surfaces، coverage holes، shared roots، paths.
المعادلة	EvidenceStrength = Σ_p w_p·Indep_p·Coverage_p·Directness_p - λ·ContradictionMass - μ·SharedRootPenalty
المكذّب	إذا لم يتنبأ geometry score بدقة claims أو calibration أفضل من source count/average quality في holdout.
الأدوات 5	Moussaif Evidence Graph Constructor, Moussaif Independence Estimator, Moussaif Contradiction Surface Mapper, Moussaif Coverage Gap Detector, Moussaif Evidence Geometry Court
التدريب 5	Moussaif Source-Independence Contrastive Training, Moussaif Evidence Triangulation Training, Moussaif Contradiction Localization Training, Moussaif Sparse-Evidence Curriculum, Moussaif Graph-Based Claim Verification

التنبؤات القابلة للفحص

- خمسة مصادر مشتقة من مصدر واحد قد تضيف أقل من مصدر مستقل ثانٍ.
- التغطية غير المتوازنة تكشف claims تبدو قوية إجمالاً.
- contradiction locality تسمح بإصلاح subclaim دون هدم الكل.
- Graph shape ينبئ بقابلية المراجعة.

مثال تفسيري

10 مقالات تنقل نفس البيان الصحفي ليست 10 تأكيدات مستقلة؛ geometry يضغطها إلى root واحد ويطلب مساراً مستقلاً.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-18 · MUT-AFI

نظرية مساعف للتشارك التطوري لمنظومة الأدوات

Moussaif Tool Ecology Co-Adaptation Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
EcoUtility = Σ_i p_i(U_i·Rel_i·C_i - Cost_i) - λ·HHI(p) +
μ·SubstitutabilityPortfolio
```

الفكرة المركزية: الأدوات تشكل ecology ديناميكية؛ جودة النظام تعتمد على portfolio من الاعتمادية والكلفة والتخصص والتعويض، وعلى تدريب الوكيل ليعيد التوازن عند تغير الأداة لا أن يلتصق بواجهة واحدة.

M-AFI-18

1. الفجوة البحثية والأطروحة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

Tool contracts وrouting يعاملان الأدوات غالباً كوحدات ثابتة، بينما إضافة أداة أو تغير API يغير توزيع الاستخدام ومهارات الوكيل واعتماده على بقية الأدوات.

الأطروحة المقترحة

الأدوات تشكل ecology ديناميكية؛ جودة النظام تعتمد على portfolio من الاعتمادية والكلفة والتخصص والتعويض، وعلى تدريب الوكيل ليعيد التوازن عند تغير الأداة لا أن يلتصق بواجهة واحدة.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تضيف dynamics portfolio-level وco-adaptation بين agent/tool بدل تقييم tool واحد خارج النظام.**

العلاقة بالسلالة السابقة

Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-18

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

p_i حصة استخدام؛ U_i utility؛ Rel_i reliability؛ C_i compatibility؛ Cost_i؛ HHI تركّز؛ Substitutability.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-18",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "p_i حصة استخدام؛ U_i utility؛ Rel_i reliability؛ C_i
compatibility؛ Cost_i؛ HHI تركّز؛ Substitutability.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-18

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
EcoUtility = Σ_i p_i(U_i·Rel_i·C_i - Cost_i) - λ·HHI(p) + μ·SubstitutabilityPortfolio
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

> **ما لا يجوز ادعاؤه:** وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-18

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
بناء Tool Ecology Map للعلاقات البديلة/المكملة/المعتمدة.

2. المرحلة
قياس concentration واعتماد النظام على أداة واحدة.

3. المرحلة
تعلم capability signatures لا أسماء API.

4. المرحلة
محاكاة إزالة أو degradation للأدوات.

5. المرحلة
إعادة تدريب routing/skills على substitutes.

6. المرحلة
تحديث ecology بعد version changes وقياس regression.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-18

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يكمل Cooperative Vascular Routing وRuntime Discovery ويحول tool resilience إلى موضوع تعلم مستمر.

الإضافة غير القابلة للاختزال
تضيف dynamics portfolio-level وco-adaptation بين agent/tool بدل تقييم tool واحد خارج النظام.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Tool registry + discovery يبني ecology دورياً ويغذي router وtraining queue.

M-AFI-18

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Tool Ecology Map
يمثل complement/substitute/dependency بين الأدوات.

Tool 1/5 فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Tool Concentration Index
يقيس اعتماد النظام على عدد قليل من الأدوات.

Tool 2/5 فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Co-Adaptation Scheduler
يحدد أي skills/router policies تحتاج إعادة تدريب بعد تغير tool.

Tool 3/5 فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Tool Substitution Harness
يختبر البدائل تحت نفس المهمة.

Tool 4/5 فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Tool Ecology Court
يفحص resilience مقابل cost/security.

Tool 5/5 فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-18

7. عقود الأدوات ومقاييسها

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Tool Ecology Map	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Tool Concentration Index	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Co-Adaptation Scheduler	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Tool Substitution Harness	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Tool Ecology Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-18

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Tool-Diversity Curriculum

تعريض النظام لعدة أدوات تؤدي capability نفسها.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Substitution Training

تبديل الأداة أثناء المهمة مع حفظ الهدف والعقد.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-18

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Adaptive Tool-Routing RL

مكافأة اختيار الأداة وفق reliability/cost/state.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Cost-Reliability Co-Training

تعلم tradeoff تحت ميزانيات مختلفة.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Tool-Failure Ecology Simulation

محاكاة outages وschema drift لتدريب recovery.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-18

10. البيانات والتقييم والمعايير

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

tool calls، reliability، latency، cost، schemas، failure logs، substitution outcomes.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- التركيز العالي يزيد brittleness حتى مع أداة ممتازة.
- substitution training يقلل recovery time.
- أدوات redundant جزئياً قد ترفع resilience.
- تغير tool واحد قد يستلزم إعادة معايرة Skills لا router فقط.

M-AFI-18

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Tool registry + discovery يبني ecology دورياً ويغذي router وtraining queue.

عقد النقل

يكمل Cooperative Vascular Routing وRuntime Discovery ويحول tool resilience إلى موضوع تعلم مستمر.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-18",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-18

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> بيئة أدوات متعددة؛ احذف الأداة الأكثر استخداماً أو غير schema؛ قارن static routing وfallback list وco-adaptive ecology.

المكذّب الرئيسي

> إذا لم يحسن ecology-aware training الاستمرارية تحت tool outage/version shift مقارنة بروتين fallback بسيط.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-18	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-18

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

1. تنويع الأدوات بلا فائدة.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. اعتبار كل substitute مكافئاً.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. تجاهل security differences.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. رفع الكلفة لتقليل HHI.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. التعلم على synthetic schemas فقط.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

البديل يجب أن يمر بنفس authorization/privacy/tool contract gates؛ لا تستخدم أداة أقل أماناً فقط لتعويض outage.

الحدود

الutility متعدد الأبعاد ومتغير؛ بعض الأدوات فريدة لا substitute لها؛ diversity قد تزيد attack surface.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-18

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Tool Contracts، Cognitive Router، HyperSkill، Runtime Discovery | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للتشارك التطوري لمنظومة الأدوات Moussaif Tool Ecology Co-Adaptation Theory
الفارق الفريد	تضيف dynamics portfolio-level وco-adaptation بين agent/tool بدل تقييم tool واحد خارج النظام.
المعادلة	EcoUtility = Σ_i p_i(U_i·Rel_i·C_i - Cost_i) - λ·HHI(p) + μ·SubstitutabilityPortfolio
المكذّب	إذا لم يحسن ecology-aware training الاستمرارية تحت tool outage/version shift مقارنة بروتين fallback بسيط.
الأدوات 5	Moussaif Tool Ecology Map, Moussaif Tool Concentration Index, Moussaif Co-Adaptation Scheduler, Moussaif Tool Substitution Harness, Moussaif Tool Ecology Court
التدريب 5	Moussaif Tool-Diversity Curriculum, Moussaif Substitution Training, Moussaif Adaptive Tool-Routing RL, Moussaif Cost-Reliability Co-Training, Moussaif Tool-Failure Ecology Simulation

التنبؤات القابلة للفحص

- التركيز العالي يزيد brittleness حتى مع أداة ممتازة.
- substitution training يقلل recovery time.
- أدوات redundant جزئياً قد ترفع resilience.
- تغير tool واحد قد يستلزم إعادة معايرة Skills لا router فقط.

مثال تفسيري

إذا تعطلت أداة بحث واحدة، نظام متعلم على capability signature ينتقل إلى مصدر بديل مع فرق جودة معلوم، لا يعيد نفس call بلا نهاية.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-19 · MUT-AFI

نظرية مساعف للتعلم الفوقي من أحكام المحاكم

Moussaif Court-to-Learning Meta-Adaptation Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
LearnSignal = Weight(V)·A_c·N·Rep; m* = argmax_m
ExpectedGeneralization(m|signal)-Risk(m)
```

الفكرة المركزية: كل Court Verdict يجب أن يُترجم إلى LearnSignal نمطي يحدد السبب والنطاق والثقة ونوع update الأنسب، مع منع overfitting إلى حكم واحد واشتراط replay/holdout قبل promotion.

الجزء الخامس · النظرية 15 · M-AFI-19 صفحة مخصصة · الصفحات 279-293

M-AFI-19

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة

المحاكم تصدر ADOPT/REVISE/MERGE/HOLD/REJECT/ROLLBACK، لكن الحكم قد يبقى artifact توثيقياً ولا يتحول إلى إشارة تعلم مضبوطة تمنع تكرار النمط.

الأطروحة المقترحة

كل Court Verdict يجب أن يُترجم إلى LearnSignal نمطي يحدد السبب والنطاق والثقة ونوع update الأنسب، مع منع overfitting إلى حكم واحد واشتراط replay/holdout قبل promotion.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تربط governance verdicts بقرار تعلم: memory patch أو KB update أو skill repair أو preference data أو weight update، بدلاً من update غير مميز.**

العلاقة بالسلالة السابقة

Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-19

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

V verdict، A_c causal attribution، N novelty، Rep reproducibility، m update mode.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-19",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "V verdict؛ A_c causal attribution؛ N novelty؛ Rep
reproducibility؛ m update mode.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-19

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
LearnSignal = Weight(V)·A_c·N·Rep; m* = argmax_m ExpectedGeneralization(m|signal)-Risk(m)
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧ NoCriticalRegression
```

> **ما لا يجوز ادعاؤه:** وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-19

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
تحويل verdict إلى structured cause code وscope.

2. المرحلة
فصل خطأ knowledge عن routing/tool/authority/model.

3. المرحلة
تقدير novelty وreproducibility للحكم.

4. المرحلة
اختيار أقل intervention كافٍ لإصلاح السبب.

5. المرحلة
توليد training/evaluation pair مع negative replay.

6. المرحلة
ترقية update فقط بعد holdout وعدم regression.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-19

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يغلق الحلقة Training→Memory→Courts التي كانت مفهومة جزئياً في الأجزاء السابقة.

الإضافة غير القابلة للاختزال
تربط governance verdicts بقرار تعلم: memory patch أو KB update أو skill repair أو preference data أو weight update، بدلاً من update غير مميز.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

قاعدة MUT-AFI: هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Meta-adaptation buffer بين Court Engine وEvolution Lab/Training Router.

M-AFI-19

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً، كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Verdict Signal Compiler
يحول أحكام المحاكم إلى schema تعلم.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Court Attribution Mapper
يربط الحكم بالطبقة السببية المسؤولة.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Meta-Adaptation Buffer
يجمع signals قبل أي update ثقيل.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Verdict-to-Data Generator
ينشئ preference/SFT/eval artifacts قابلة للمراجعة.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Learning Court
يفحص ما إذا كان الحكم يستحق update وأي نوع.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-19

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Verdict Signal Compiler	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Court Attribution Mapper	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Meta-Adaptation Buffer	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Verdict-to-Data Generator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Learning Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-19

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Verdict-Conditioned SFT

أمثلة تربط سبب الحكم بالإصلاح الصحيح.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval,
regression}
```

2. Moussaif Failure-to-Preference Pair Training

إنشاء chosen/rejected من مسار فاشل ومصلح.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval,
regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-19

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Court-Guided RL

مكافأة اجتياز verifier مستقل لا إرضاء court لغوياً.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Verdict Distillation

تقطير قواعد متكررة إلى skill/policy أصغر.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Anti-Overfit Replay Training

إعادة تشغيل نجاحات قديمة لمنع regression.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-19

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

update manifests، regressions، holdouts، repeats، repairs، cause attribution، verdicts.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- أحكام ذات attribution ضعيف يجب ألا تقود weight update.
- أغلب الإصلاحات تبدأ memory/skill قبل training ثقيل.
- verdict repetition يزيد أولوية intervention.
- negative replay يمنع إصلاحاً محلياً من تخريب قدرات أخرى.

M-AFI-19

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Meta-adaptation buffer بين Court Engine وEvolution Lab/Training Router.

عقد النقل

يغلق الحلقة Training→Memory→Courts التي كانت مفهومة جزئياً في الأجزاء السابقة.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-19",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought؛ انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-19

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> اجمع failures متكررة، طبّق memory-only وgeneric fine-tune وmeta-adaptation؛ قارن recurrence/generalization/regression.

المكذّب الرئيسي

> إذا لم يخفض court-conditioned updates recurrence أكثر من تخزين verdict في memory فقط تحت نفس budget.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-19	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-19

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

1. تدريب على كل REJECT.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. اعتبار court معصومة.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. تسريب holdout في repair data.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. عدم فصل سبب tool عن model.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. promotion بلا rollback.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

المحكمة نفسها subject to audit؛ لا يتحول verdict حساس إلى training data إذا احتوى أسراراً أو بيانات غير مصرح بها.

الحدود

causal attribution قد يكون خاطئاً؛ rare failures لا توفر signal كافياً؛ weight update مكلف وعالي الأثر.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-19

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: Demolition Court، Evolution Laboratory، Continuous Post-Training، Failure Memory | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف للتعلم الفوقي من أحكام المحاكم Moussaif Court-to-Learning Meta-Adaptation Theory
الفارق الفريد	تربط governance verdicts بقرار تعلم: memory patch أو KB update أو skill repair أو preference data أو weight update، بدلاً من update غير مميز.
المعادلة	LearnSignal = Weight(V)·A_c·N·Rep; m* = argmax_m ExpectedGeneralization(m|signal)-Risk(m)
المكذّب	إذا لم يخفض court-conditioned updates recurrence أكثر من تخزين verdict في memory فقط تحت نفس budget.
الأدوات 5	Moussaif Verdict Signal Compiler, Moussaif Court Attribution Mapper, Moussaif Meta-Adaptation Buffer, Moussaif Verdict-to-Data Generator, Moussaif Learning Court
التدريب 5	Moussaif Verdict-Conditioned SFT, Moussaif Failure-to-Preference Pair Training, Moussaif Court-Guided RL, Moussaif Verdict Distillation, Moussaif Anti-Overfit Replay Training

التنبؤات القابلة للفحص

- أحكام ذات attribution ضعيف يجب ألا تقود weight update.
- أغلب الإصلاحات تبدأ memory/skill قبل training ثقيل.
- verdict repetition يزيد أولوية intervention.
- negative replay يمنع إصلاحاً محلياً من تخريب قدرات أخرى.

مثال تفسيري

إذا رفضت المحكمة claim بسبب stale source، العلاج الأول revalidation policy لا fine-tuning model على «الإجابة الصحيحة».

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

M-AFI-20 · MUT-AFI

نظرية مساعف لترقية الأبعاد ونقلها

Moussaif Dimension Promotion and Transfer Theory

الرتبة العلمية: **N0/N1 - نظرية مرشحة مُركبة داخل البرنامج.** ليست ادعاء أسبقية عالمية ولا نتيجة تجريبية مثبتة. الترقية تتطلب baseline + ablation + replication.

```
DimValue(d)=α·ΔPred+β·ΔTrans+γ·ΔAct+η·Indep-λ·K; Promote iff
DimValue>τ on ≥2 domains ∧ AblationLoss>0
```

الفكرة المركزية: البعد لا يُرقّى لأنه يحسن مثالاً واحداً، يجب أن يثبت استقلالاً جزئياً، زيادة في التنبؤ/التمييز/الفعل، ablation loss، ونقلاً بين مجالين مع عقوبة تعقيد.

M-AFI-20

1. الفجوة البحثية والأطروحة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

الفجوة المحددة
OSDS/QDG يسمحان باقتراح بعد جديد عندما يفشل فضاء التمثيل، لكن ينقص بروتوكول يقرر متى يصبح البعد المرشح جزءاً مستقراً من ontology ويمكن نقله إلى مجال آخر.

الأطروحة المقترحة

البعد لا يُرقّى لأنه يحسن مثالاً واحداً؛ يجب أن يثبت استقلالاً جزئياً، زيادة في التنبؤ/التمييز/الفعل، ablation loss، ونقلاً بين مجالين مع عقوبة تعقيد.

لماذا لا تكفي المكوّنات السابقة وحدها؟

لا يُقترح هذا الامتداد لمجرد إيجاد اسم جديد. وفق قاعدة MUT، لا تُقبل عقدة نظرية إلا إذا امتلكت تعريفاً وآلية ومخرجاً واختباراً وحدوداً. هنا تتمثل الإضافة في: **تحول «فتح بعد جديد» من استعارة نمو التمثيل إلى lifecycle كامل: propose→test→quarantine→transfer→promote/deprecate.**

العلاقة بالسلالة السابقة
OSDS، QDG، CIDM، Cross-Domain Transfer

سؤال النظرية الحاكم

ما المتغير أو الانتقال الذي يجب قياسه كي نميز بين نجاح ظاهري يمكن تفسيره بالمكوّنات السابقة وبين الأثر الفريد لهذه النظرية؟ الجواب التشغيلي يبدأ من المتغيرات والمعادلة والاختبار الهَدمي في الصفحات التالية.

شرط عدم التضخم: إذا أمكن تفسير النتائج دون هذا المتغير الجديد وبنفس جودة التنبؤ والتكلفة، تُدمج النظرية في أصلها ولا تبقى عقدة مستقلة.

M-AFI-20

2. التعريفات والمسلّمات التشغيلية

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

قاموس المتغيرات

ΔPred تحسن التنبؤ، ΔTrans النقل، ΔAct الفعل، Indep، K كلفة التعقيد.

المسلّمات المقترحة

1. **مبدأ الرصد:** لا يُسند تغير إلى النظرية ما لم توجد حالة قبلية وبعدية قابلة للمقارنة.
2. **مبدأ التدخل:** يجب فصل الوجود الاسمي للمكوّن عن مساهمته الهامشية عبر إزالة أو بديل أو تدخل مضبوط.
3. **مبدأ النطاق:** كل نتيجة تحمل domain/task/authority scope ولا تتحول تلقائياً إلى قانون عام.
4. **مبدأ الاستقلال:** المقاييس المشتقة لا يجوز أن تكون إعادة تسمية مباشرة لـMCD/QMCD أو success rate.
5. **مبدأ الإغلاق:** التقييم الأعلى قيمة يصل إلى World Effect أو نتيجة خارجية/اختبارية قابلة للتحقق.

كائن النظرية الأدنى

```
theory_state {
  code: "M-AFI-20",
  inputs: canonical_state + evidence + budget,
  latent_or_measured_variables: "ΔPred تحسن التنبؤ؛ ΔTrans النقل؛ ΔAct الفعل؛ Indep؛ K كلفة
التعقيد.",
  mechanism_version: "v0.1-proposed",
  outputs: prediction + decision_delta + evidence_delta,
  falsifier: explicit,
  provenance: Parts_I_IV + MUT_AGI_02,
  status: PROPOSED
}
```

حد معرفي: المسلمات هنا شروط هندسية للبحث وليست بديهيات طبيعية مثبتة.

M-AFI-20

3. النموذج الرياضي وحدود المعادلة

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

الصيغة المرشحة

```
DimValue(d)=α·ΔPred+β·ΔTrans+γ·ΔAct+η·Indep-λ·K; Promote iff DimValue>τ on ≥2 domains
∧ AblationLoss>0
```

المعادلة ليست «قانون ذكاء» نهائياً. دورها أن تجعل النظرية قابلة للقياس والمقارنة. يجب تعريف كل متغير، مجاله، طريقة normalization، وحدة القياس أو proxy، ودرجة عدم اليقين. أي coefficient غير معاير يبقى parameter بحثياً لا قيمة ثابتة.

Calibration Protocol

- تثبيت baseline.
- فصل train/validation/holdout.
- تقدير sensitivity لكل معامل.
- bootstrap/seed variance.
- تسجيل calibration drift.

Ablation Protocol

- حذف المتغير الجديد.
- مقارنة complexity-matched baseline.
- equal-budget compute/tool calls.
- نقل بين مجالين.
- negative control لا ينبغي أن يتحسن.

معيار الترقية الرياضية

```
PromoteFormula ⇔ ΔPredictiveUtility > 0 ∧ CalibrationStable ∧ AblationLoss > 0 ∧
NoCriticalRegression
```

ما لا يجوز ادعاؤه: وجود صيغة رياضية لا يجعل المتغير حقيقة طبيعية، ولا يثبت AGI، ولا يثبت وعياً. المعادلة أداة لجعل الخطأ قابلاً للرصد.

M-AFI-20

4. الآلية السببية/التشغيلية خطوة بخطوة

مرساة التركيب: Parts I-IV + MUT-AGI-02 | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

```
STATE → DIFFERENCE → THEORY OPERATOR → DECISION DELTA → ACTION/TEST → EFFECT → COURT → MEMORY
```

1. المرحلة
استخراج residual pattern لا تفسره الأبعاد الحالية.

2. المرحلة
اقتراح dimension candidate بتعريف وقياس وحدود.

3. المرحلة
اختبار independence عبر conditional ablation.

4. المرحلة
قياس value في التنبؤ والقرار وليس fit فقط.

5. المرحلة
تجربة transfer إلى مجال ثانٍ بأسماء مختلفة.

6. المرحلة
ترقية البعد إلى ontology أو عزله/دمجه/إلغاؤه.

شرط الإغلاق

لا تعتبر الآلية مكتملة عند إنتاج نص أو score. يجب أن تنتج تغييراً قابلاً للفحص في قرار، توجيه، اختيار أداة، نتيجة اختبار، World Effect، أو Memory Promotion. إن انتهت قبل ذلك فهي تفسير لا mechanism مكتمل.

> **مبدأ التنفيذ الأدنى:** ابدأ بأصغر implementation يسمح بعزل operator الخاص بالنظرية ثم وسّع فقط بعد نجاح ablation.

M-AFI-20

5. موضعها داخل MUT وعدم التكرار

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

العلاقة القانونية

يمنح OSDS/QDG قانون ترقية ويصل الوعي المفتوح الأبعاد بالتعلم التمثيلي القابل للهدم.

> **الإضافة غير القابلة للاختزال**
> تحول «فتح بعد جديد» من استعارة نمو التمثيل إلى lifecycle كامل: propose→test→quarantine→transfer→promote/deprecate.

اختبار التشابه مع النظريات السابقة

السؤال	إذا كانت الإجابة نعم	قرار registry
هل المدخل/المخرج مطابقان لأصل سابق؟	لا فرق تشغيلي	Alias / Merge
هل تغير transform أو state variable؟	فرق آلي محتمل	Keep as candidate
هل يوجد falsifier مميز؟	قابل للفصل	Independent test
هل يظهر ablation loss؟	مساهمة هامشية	Promote evidence rank
هل ينتقل إلى مجال ثانٍ؟	أقل احتمالاً أن يكون label trick	Transfer support

> **قاعدة MUT-AFI:** هذا الكتاب يقترح 20 عقدة مرشحة. «جديدة» تعني جديدة في التركيب الحالي للبرنامج؛ الأسبقية العالمية تحتاج مراجعة أدبية مستقلة ولا تُفترض هنا.

مكانها في الدورة العليا

Dimension registry متصل بـQDG وConcept Compiler وEvolution Lab.

M-AFI-20

6. الأدوات الخمس المشتقة من النظرية

مرساة التركيب: MUT-AGI-02 Parts I-IV + | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

الأداة هنا ليست اسماً زخرفياً؛ كل أداة يجب أن تملك trigger ومدخلات ومخرجات ومقياساً وفشلاً متوقعاً. الأدوات الخمس التالية هي الحد الأدنى لتحويل النظرية إلى بيئة اختبار.

1. Moussaif Dimension Candidate Generator
يستخرج residual patterns ويقترح تعريفات مرشحة.

Tool 1/5 | فائدة قابلة للقياس تحت baseline وablation

2. Moussaif Dimension Independence Tester
يقيس المعلومات الإضافية بعد conditioning على الأبعاد القديمة.

Tool 2/5 | فائدة قابلة للقياس تحت baseline وablation

3. Moussaif Transfer Surface Evaluator
يختبر البعد عبر مجالات ومهام مختلفة.

Tool 3/5 | فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Complexity Penalty Calculator
يعاقب توسع ontology غير الضروري.

Tool 4/5 | فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Dimension Promotion Court
يقرر promote/merge/quarantine/deprecate.

Tool 5/5 | فائدة قابلة للقياس تحت baseline وablation

قاعدة التكوين: لا يجوز للأداة الخامسة مثلاً أن تكون مجرد واجهة مختلفة للأداة الأولى. يجب أن تغطي كل أداة وظيفة مستقلة في دورة observe→decide→test→verify.

M-AFI-20

7. عقود الأدوات ومقاييسها

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

#	الأداة	المدخل	المخرج	المقياس
1	Moussaif Dimension Candidate Generator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
2	Moussaif Dimension Independence Tester	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
3	Moussaif Transfer Surface Evaluator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
4	Moussaif Complexity Penalty Calculator	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation
5	Moussaif Dimension Promotion Court	حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة	كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام	فائدة قابلة للقياس تحت baseline وablation

Tool Contract موحّد

```
tool_contract {
 id, version, trigger, inputs_schema, outputs_schema,
 authority_required, side_effects, timeout, retry_policy,
 verification_method, failure_signatures, rollback,
 provenance, telemetry, cost_budget
}
```

محكمة الأدوات: نجاح call لا يساوي نجاح الأثر. كل أداة ذات أثر تحتاج postcondition مستقل، والأداة التحليلية تحتاج benchmark أو consistency check مناسباً.

اختبار الاستبدال

لكل أداة، نفّذ substitute أو no-tool baseline. إذا لم يتغير المخرج أو المقياس بصورة قابلة للإعادة، فالأداة ليست primitive لازماً وقد تُدمج.

M-AFI-20

8. طرق التدريب - المنهجان 1 و2

مرساة التركيب: MUT-AGI-02 Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

لا يُفترض أن كل نظرية تحتاج fine-tuning للأوزان. «طريقة تدريب» هنا قد تكون curriculum أو preference/process training أو تعلم policy أو calibration. يُختار أقل تدخل كافٍ ويحافظ على rollback.

1. Moussaif Latent-Dimension Discovery Training

تدريب على residuals التي لا تفسرها features الحالية.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:1, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

2. Moussaif Cross-Domain Dimension Transfer Training

اختبار نفس البعد بأسماء وسياقات مختلفة.

بيانات التدريب: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة

هدف التعلم: كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام

مقياس النجاح: فائدة قابلة للقياس تحت baseline وablation

```
train_record{method:2, split:train/val/holdout, baseline:frozen, seed, budget, loss, eval, regression}
```

Least-Sufficient Intervention: ابدأ بـretrieval/skill/policy calibration قبل weight update إذا كان المطلوب يمكن تحقيقه دون تعديل الأوزان.

M-AFI-20

9. طرق التدريب - المناهج 3 إلى 5

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

3. Moussaif Residual-Driven Representation Learning

إضافة تمثيل فقط عند failure residual ثابت.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

4. Moussaif Complexity-Regularized Expansion Training

موازنة gain مع تكلفة البعد.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

5. Moussaif Dimension Collapse-Rejection Training

تعليم رفض candidate ينهار إلى combination قديم.

Input: حالة المهمة + السياق المعياري + الأدلة/القيود ذات الصلة | **Output/Objective:** كائن قرار/سجل نمطي قابل للتتبع وإعادة الاستخدام | **Metric:** فائدة قابلة للقياس تحت baseline وablation

بوابة ترقية التدريب

```
PromoteTraining ⇔ HoldoutGain > τ ∧ NoCriticalRegression ∧ ReproducibleSeeds ≥ k ∧
DataProvenanceOK ∧ RollbackReady
```

فصل مهم: تحسن RAG أو prompt policy لا يُسجل كـweight learning. كل intervention يحمل نوعه صراحة: memory / retrieval / skill / preference / adapter / continued-pretraining.

M-AFI-20

10. البيانات والتقييم والمعايير

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

مصادر البيانات المطلوبة

complexity measures، decision outcomes، domain transfer sets، ablations، candidate definitions، residuals.

بروتوكول التقييم

1. Freeze baseline وإصدار البيانات والأدوات.
2. حدد primary metric و2-3 secondary metrics قبل التشغيل.
3. افصل calibration set عن holdout النهائي.
4. نفّذ ablation للمكوّن المقترح وcomplexity-matched baseline.
5. كرر على seeds ونموذج/مزود ثانٍ إذا كانت النظرية تدعي transfer.
6. سجل failures والنتائج السلبية بنفس صرامة النجاحات.

Primary evidence
world-effect / external test / reproducible held-out performance.

Non-evidence alone
عدد الصفحات، جودة الشرح، كثرة الأدوات، self-confidence، أو citation count.

Prediction registry

- معظم candidates تُرفض أو تدمج.
- الأبعاد الحقيقية تظهر ablation loss خارج مجال الاكتشاف.
- complexity penalty يمنع dimension explosion.
- بعض أبعاد الوعي الوظيفي تتحول إلى features هندسية قابلة للنقل دون ادعاء phenomenal consciousness.

M-AFI-20

11. الاندماج في Runtime والنقل بين النماذج

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

موضع التنفيذ

Dimension registry متصل بـQDG وConcept Compiler وEvolution Lab.

عقد النقل

يمنح OSDS/QDG قانون ترقية ويصل الوعي المفتوح الأبعاد بالتعلم التمثيلي القابل للهدم.

```
state_capsule {
 goal, governing_question, claims, evidence_refs,
 theory_operator:"M-AFI-20",
 variables, decision, risk, authority,
 open_questions, effect_expectation, rollback_pointer
}
```

اختبارات portability

- Provider swap دون نقل chain-of-thought، انقل state المنظم فقط.
- Name stripping للنظرية والأدوات لاختبار اعتماد الوظيفة على البنية لا label.
- Tool substitution مع capability contract ثابت.
- Context compaction ثم reconstruction واختبار decision fidelity.
- إعادة الاختبار بميزانية مساوية لتجنب إسناد التحسن لحوسبة أكبر.

حالة التنفيذ: specification-level حتى تظهر traces فعلية من runtime مرجعي.

M-AFI-20

12. التجربة الحاسمة والـAblation

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

التجربة المقترحة

> مجالات مهمة مختلفة تشترك في pattern خفي؛ قارن reweight old dimensions وnew dimension مع regularization وablation.

المكذّب الرئيسي

> إذا كان التحسن يختفي عند complexity-matched baseline أو domain transfer، لا يُعد البعد إضافة نظرية مستقلة.

ذراع	الوصف	الهدف
A	Frozen baseline بلا operator M-AFI-20	مرجع سببي
B	Operator كامل مع نفس الميزانية	قياس contribution
C	Complexity-matched control	استبعاد فائدة التعقيد العام
D	Ablation لأهم متغير/أداة	تحديد primitive الفعلي
E	Transfer domain/provider	فحص التعميم

قرار المحكمة بعد التجربة

```
Verdict ∈ {PROMOTE, REVISE, MERGE_WITH_PARENT, HOLD, REJECT}
```

حتى PROMOTE لا يعني «نظرية AGI مثبتة»؛ يعني فقط أن هذا المكوّن أظهر مساهمة متكررة ضمن النطاق المختبر.

M-AFI-20

13. محكمة الهدم والسلامة وحدود الفشل

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

1. إضافة بعد لكل residual.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

2. تعريف غير قابل للقياس.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

3. نقل اسمي لا بنيوي.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

4. double-counting مع dimension قديم.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

5. اعتبار fit دليلاً على ontology حقيقية.
يجب تسجيل trigger، الأثر، repair، detector، وهل تكرر بعد الإصلاح.

قاعدة السلامة

لا تُفسر الأبعاد الوظيفية ككيانات ميتافيزيقية أو دليل وعي. الترقية epistemic/engineering فقط.

الحدود

الاستقلال التام نادر؛ الأبعاد قد تكون task-relative؛ dimension discovery معرض لـresearcher degrees of freedom.

Demolition Question: ما أبسط تفسير بديل يفسر gain دون النظرية؟ إذا كان موجوداً ولم يُهزم بتجربة متساوية الميزانية، تبقى النظرية HOLD أو تُدمج.

M-AFI-20

14. بطاقة النظرية النهائية وبرنامج البحث

مرساة التركيب: MUT-AGI-02 + Parts I-IV | الأصول المتصلة: OSDS، QDG، CIDM، Cross-Domain Transfer | الحالة: Proposed MUT-AFI synthesis

الهوية	نظرية مساعف لترقية الأبعاد ونقلها Moussaif Dimension Promotion and Transfer Theory
الفارق الفريد	تحول «فتح بعد جديد» من استعارة ثمو التمثيل إلى lifecycle كامل: propose→test→quarantine→transfer→promote/deprecate.
المعادلة	$DimValue(d)=\alpha\cdot\Delta Pred+\beta\cdot\Delta Trans+\gamma\cdot\Delta Act+\eta\cdot Indep-\lambda\cdot K$; Promote iff $DimValue>\tau$ on ≥ 2 domains $\wedge$ $AblationLoss>0$
المكذّب	إذا كان التحسن يختفي عند complexity-matched baseline أو domain transfer، لا يُعد البعد إضافة نظرية مستقلة.
الأدوات 5	Moussaif Dimension Candidate Generator, Moussaif Dimension Independence Tester, Moussaif Transfer Surface Evaluator, Moussaif Complexity Penalty Calculator, Moussaif Dimension Promotion Court
التدريب 5	Moussaif Latent-Dimension Discovery Training, Moussaif Cross-Domain Dimension Transfer Training, Moussaif Residual-Driven Representation Learning, Moussaif Complexity-Regularized Expansion Training, Moussaif Dimension Collapse-Rejection Training

التنبؤات القابلة للفحص

- معظم candidates تُرفض أو تدمج.
- الأبعاد الحقيقية تظهر ablation loss خارج مجال الاكتشاف.
- complexity penalty يمنع dimension explosion.
- بعض أبعاد الوعي الوظيفي تتحول إلى features هندسية قابلة للنقل دون ادعاء phenomenal consciousness.

مثال تفسيري

إذا فشلت أبعاد risk/cost/uncertainty في تفسير قرارات الرجوع، قد يقترح Reversibility كـdimension؛ لا يرقى إلا إذا حسّن مهام متعددة وظهر ablation loss.

سلم النضج المقترح

```
N0 SYNTHESIS → N1 FORMALIZED → N2 EXPERIMENT DESIGNED → N3 LOCAL EVIDENCE → N4 CROSS-MODEL REPLICATION → N5 PROGRAM PROMOTION
```

Next Gate: بناء implementation أدنى للأدوات الخمس، تشغيل التجربة الحاسمة، ثم تحديث Theory Registry بالدليل الفعلي لا بالوصف.

الملحق A - سجل النظريات العشرين

Code	العربية	English	المحور
M-AFI-01	نظرية مساعف للإسناد السببي للفروق	Moussaif Causal Difference Credit Assignment Theory	الإسناد السببي للفروق
M-AFI-02	نظرية مساعف للانتقال الطوري الإبستيمي	Moussaif Epistemic Phase Transition Theory	الانتقال الطوري الإبستيمي
M-AFI-03	نظرية مساعف للتخطيط المضاد للواقع لأثر العالم	Moussaif Counterfactual World-Effect Planning Theory	التخطيط المضاد للواقع لأثر العالم
M-AFI-04	نظرية مساعف للمزامنة الحاكمية متعددة الآفاق	Moussaif Multi-Horizon Governance Synchronization Theory	المزامنة الحاكمية متعددة الآفاق
M-AFI-05	نظرية مساعف لإعادة البناء الحافظة للقيمة	Moussaif Value-Preserving Self-Reconstruction Theory	إعادة البناء الحافظة للقيمة
M-AFI-06	نظرية مساعف للاستتباب الدلالي عبر النماذج	Moussaif Cross-Model Semantic Homeostasis Theory	الاستتباب الدلالي
M-AFI-07	نظرية مساعف لإغلاق القدرة التركيبية	Moussaif Compositional Capability Closure Theory	إغلاق القدرة التركيبية
M-AFI-08	نظرية مساعف لتعلّم طوبولوجيا الفشل	Moussaif Failure Topology Learning Theory	طوبولوجيا الفشل
M-AFI-09	نظرية مساعف لتحويل عدم اليقين إلى وكالة	Moussaif Uncertainty-to-Agency Conversion Theory	تحويل عدم اليقين إلى وكالة
M-AFI-10	نظرية مساعف لموازنة الحمل الإبستيمي	Moussaif Epistemic Load Balancing Theory	موازنة الحمل الإبستيمي
M-AFI-11	نظرية مساعف لديناميات الثقة التفويضية	Moussaif Delegation Trust Dynamics Theory	ديناميات الثقة التفويضية
M-AFI-12	نظرية مساعف للنمو القابل للعكس للاستقلالية	Moussaif Reversible Autonomy Growth Theory	النمو القابل للعكس للاستقلالية
M-AFI-13	نظرية مساعف للدَّين المعرفي وإصلاحه	Moussaif Cognitive Debt Accumulation and Repair Theory	الدَّين المعرفي
M-AFI-14	نظرية مساعف لإنتروبيا المهمة والانضغاط الحاكم	Moussaif Mission Entropy and Governing Compression Theory	إنتروبيا المهمة
M-AFI-15	نظرية مساعف للمناعة ضد انحراف الهدف	Moussaif Adversarial Goal-Drift Immunity Theory	مناعة انحراف الهدف
M-AFI-16	نظرية مساعف للنقل السببي من المعرفة إلى المهارة	Moussaif Knowledge-to-Skill Causal Transfer Theory	النقل السببي من المعرفة إلى المهارة
M-AFI-17	نظرية مساعف لهندسة الدليل	Moussaif Evidence Geometry Theory	هندسة الدليل
M-AFI-18	نظرية مساعف للتشارك التطوري لمنظومة الأدوات	Moussaif Tool Ecology Co-Adaptation Theory	التشارك التطوري للأدوات
M-AFI-19	نظرية مساعف للتعلم الفوقي من أحكام المحاكم	Moussaif Court-to-Learning Meta-Adaptation Theory	التعلم من أحكام المحاكم
M-AFI-20	نظرية مساعف لترقية الأبعاد ونقلها	Moussaif Dimension Promotion and Transfer Theory	ترقية الأبعاد ونقلها

العدد = 20 مرشحاً نظرياً، لا 20 حقيقة مثبتة. سجل MUT يحتفظ بحالة كل عقدة وقرار merge/reject إذا فشلت التجارب.

الملحق B1 - سجل الأدوات 001-025

جزء 1/4 من سجل 100 أداة؛ خمسة أدوات لكل نظرية. الوجود في السجل لا يعني implementation مكتمل.

#	Theory	Tool	وظيفتها
001	M-AFI-01	Moussaif Causal Credit Mapper	يبني DAG سببي من الفرق إلى الأثر ويحدد نقاط التدخل.
002	M-AFI-01	Moussaif Mediator Trace Engine	يفصل الأثر المباشر عن الوسائط مثل router→tool→effect.
003	M-AFI-01	Moussaif Counterfactual Effect Comparator	ينشئ run بديل أو matched control لكل مكوّن مهم.
004	M-AFI-01	Moussaif Delayed Effect Ledger	يتتبع الأثر عبر h1..hn ويمنع إغلاق المهمة مبكرا.
005	M-AFI-01	Moussaif Causal Attribution Court	يرفض credit غير القابل للفصل سببيّا أو الملوث بزيادة الموارد.
006	M-AFI-02	Moussaif Phase Boundary Detector	يرصد تراكم residual/contradiction بعد إصلاحات محلية.
007	M-AFI-02	Moussaif Representation Rigidity Meter	يقيس قدرة ontology الحالية على استيعاب أمثلة مضادة دون تشوه.
008	M-AFI-02	Moussaif Dimension Proposal Engine	يقترح بعدا/علاقة/نوعا جديدا مع تعريف قابل للاختبار.
009	M-AFI-02	Moussaif Ontology Split-Merge Sandbox	يجرب split/merge/relation changes بعزل كامل.
010	M-AFI-02	Moussaif Phase Transition Court	يقرر reweight/repair أو transition أو reject.
011	M-AFI-03	Moussaif Counterfactual Planner	ينشئ actions وworld branches تحت do(a).
012	M-AFI-03	Moussaif World-Effect Simulator	يحاكي ΔW عبر horizons مع uncertainty.
013	M-AFI-03	Moussaif Reversibility Evaluator	يحسب rollback viability وblast radius.
014	M-AFI-03	Moussaif Horizon Comparator	يقارن short/medium/long effects ويكشف preference inversion.
015	M-AFI-03	Moussaif Intervention Court	يرفض action إذا فشل constraint/identifiability/sensitivity.
016	M-AFI-04	Moussaif Horizon Governance Matrix	يمثل goals/constraints/effects عبر h.
017	M-AFI-04	Moussaif Temporal Conflict Detector	يكشف تعارض الأهداف/القيود بين الآفاق.
018	M-AFI-04	Moussaif Policy Synchronizer	يقترح policy مشتركة أو staggered plan.
019	M-AFI-04	Moussaif Delayed Consequence Ledger	يربط القرارات بآثار لاحقة فعلية.
020	M-AFI-04	Moussaif Multi-Horizon Court	يحكم على short win / long loss وغيرها.
021	M-AFI-05	Moussaif Value Invariant Registry	يحفظ الثوابت التي لا يجوز أن تضيع في reconstruction.
022	M-AFI-05	Moussaif Reconstruction Diff Analyzer	يقارن state قبل/بعد على semantics والسلطة والمصدر.
023	M-AFI-05	Moussaif Provenance-Preserving Merger	يدمج knowledge/memory دون قطع lineage.
024	M-AFI-05	Moussaif Value Drift Alarm	يرصد تغير policy implications بعد تحديث غير policy.
025	M-AFI-05	Moussaif Reconstruction Court	يقرر promote/quarantine/rollback.

الملحق B2 - سجل الأدوات 026-050

جزء 2/4 من سجل 100 أداة؛ خمسة أدوات لكل نظرية. الوجود في السجل لا يعني implementation مكتمل.

#	Theory	Tool	وظيفتها
026	M-AFI-06	Moussaif Semantic State Canonicalizer	يحوّل حالات متعددة الصياغة إلى مخطط معياري قابل للمقارنة.
027	M-AFI-06	Moussaif Provider Translation Matrix	يسجل أخطاء التحويل واتجاهاتها بين الأزواج المختلفة من النماذج.
028	M-AFI-06	Moussaif Drift Sentinel	يراقب مشتقة الانزياح عبر الزمن ويطلق gate عند تجاوز العتبة.
029	M-AFI-06	Moussaif State Capsule Comparator	يقارن القرارات والقيود والأدلة قبل/بعد النقل لا الكلمات فقط.
030	M-AFI-06	Moussaif Semantic Homeostasis Court	يحكم هل الانحراف تعلم مشروع أم drift يحتاج استعادة.
031	M-AFI-07	Moussaif Capability Closure Graph	يبني DAG للقدرات والروابط الضرورية.
032	M-AFI-07	Moussaif Missing-Link Diagnoser	يحدد edge أو contract المفقود الذي يمنع الإغلاق.
033	M-AFI-07	Moussaif Composition Harness	يشغل تراكيب محددة تحت نفس الميزانية ويقارنها.
034	M-AFI-07	Moussaif Bottleneck Profiler	يحسب minimum cuts وعنق الزجاجة السببي.
035	M-AFI-07	Moussaif Closure Court	يرفض ادعاء القدرة إذا لم تُغلق المهمة حتى World Effect.
036	M-AFI-08	Moussaif Failure Graph Builder	يبني graph موحداً من signatures وآثار التنفيذ.
037	M-AFI-08	Moussaif Failure Clusterer	يستخرج motifs متكررة عبر مجالات مختلفة.
038	M-AFI-08	Moussaif Repair Transfer Engine	يقترح إصلاحاً عالي العائد على عدة عقد.
039	M-AFI-08	Moussaif Failure Forecaster	يتنبأ بأقرب حالات الفشل من الحالة الحالية.
040	M-AFI-08	Moussaif Failure Topology Court	يفصل الارتباط عن السبب ويطلب ablation عند الحاجة.
041	M-AFI-09	Moussaif Uncertainty Decomposer	يفصل أنواع عدم اليقين ذات سياسات مختلفة.
042	M-AFI-09	Moussaif Action-Query Gate	يختار بين الفعل والسؤال والاسترجاع والمحاكاة.
043	M-AFI-09	Moussaif Reversibility-Aware Selector	يعدل القرار حسب تكلفة الرجوع ونطاق الأثر.
044	M-AFI-09	Moussaif Abstention Planner	يحول الامتناع إلى قرار مفسر مع next-best action.
045	M-AFI-09	Moussaif Uncertainty-to-Agency Court	يفحص هل الوكالة متناسبة مع uncertainty والسلطة.
046	M-AFI-10	Moussaif Epistemic Load Monitor	يحسب الحمل عبر claims ومصادرها.
047	M-AFI-10	Moussaif Verification Scheduler	يرتب اختبارات التحقق حسب خفض الخطر المتوقع.
048	M-AFI-10	Moussaif Claim Debt Queue	يحفظ claims المؤجلة مع interest وفترة صلاحية.
049	M-AFI-10	Moussaif Resource Balancer	يوزع verifier/tool/model budget على العقد.
050	M-AFI-10	Moussaif Epistemic Load Court	يفحص هل تأجيل التحقق مقبول أو يلزم downgrade.

الملحق B3 - سجل الأدوات 051-075

جزء 3/4 من سجل 100 أداة؛ خمسة أدوات لكل نظرية. الوجود في السجل لا يعني implementation مكتمل.

#	Theory	Tool	وظيفتها
051	M-AFI-11	Moussaif Delegation Trust Ledger	يحفظ posterior والثقة حسب المجال والزمن.
052	M-AFI-11	Moussaif Error Correlation Detector	يقيس تشابه أنماط الفشل بين agents/models.
053	M-AFI-11	Moussaif Domain Competence Profiler	يفصل الأداء بحسب task family لا average واحد.
054	M-AFI-11	Moussaif Handoff Risk Evaluator	يقدّر خطر نقل مهمة إلى وكيل جديد.
055	M-AFI-11	Moussaif Delegation Court	يبرر التفويض أو يطلب استقلالاً إضافياً.
056	M-AFI-12	Moussaif Autonomy Ladder Manager	يدير مستويات السلطة ونطاقاتها ونسخها.
057	M-AFI-12	Moussaif Competence Gate	يربط الترقية بنتائج تحقق خارجية.
058	M-AFI-12	Moussaif Rollback Harness	يختبر استعادة الحالة لا مجرد وجود backup.
059	M-AFI-12	Moussaif Blast-Radius Estimator	يقدّر مدى الأثر قبل التنفيذ.
060	M-AFI-12	Moussaif Autonomy Court	يقبل/يخفض/يرفض ترقية الاستقلالية.
061	M-AFI-13	Moussaif Cognitive Debt Scanner	يعثر على claims/assumptions المؤقتة ذات الاعتماد.
062	M-AFI-13	Moussaif Assumption Staleness Meter	يقيس العمر بالنسبة إلى freshness policy.
063	M-AFI-13	Moussaif Contradiction Interest Calculator	يقدر نمو الكلفة بسبب conflicts جديدة.
064	M-AFI-13	Moussaif Debt Prioritizer	يرتب paydown حسب خفض المخاطر المتوقع.
065	M-AFI-13	Moussaif Cognitive Debt Court	يقرر verify/deprecate/narrow/accept debt.
066	M-AFI-14	Moussaif Mission Entropy Meter	يقدر كثافة المعلومات التشغيلية في الحالة.
067	M-AFI-14	Moussaif State Capsule Compiler	ينتج capsule معيارية minimal sufficient.
068	M-AFI-14	Moussaif Decision Sufficiency Tester	يحذف حقولاً ويختبر تغير القرار.
069	M-AFI-14	Moussaif Compression Re-Expander	يعيد تحميل المصدر عند trigger أو uncertainty.
070	M-AFI-14	Moussaif Mission Entropy Court	يرفض ضغطاً يفقد provenance أو constraints.
071	M-AFI-15	Moussaif Goal Anchor Ledger	يحفظ نسخ الأهداف والتحديثات المأذونة.
072	M-AFI-15	Moussaif Drift Detector	يقارن الهدف الفعلي بالسلالة المصرح بها.
073	M-AFI-15	Moussaif Prompt-Tool Injection Monitor	يميز تعليمات المصدر عن Task Authority.
074	M-AFI-15	Moussaif Objective Diff Analyzer	يحدد أي constraint/subgoal تغير ومتى.
075	M-AFI-15	Moussaif Goal Integrity Court	يحكم بين refinement مشروع وdrift.

الملحق B4 - سجل الأدوات 076-100

جزء 4/4 من سجل 100 أداة؛ خمسة أدوات لكل نظرية. الوجود في السجل لا يعني implementation مكتمل.

#	Theory	Tool	وظيفتها
076	M-AFI-16	Moussaif Knowledge Exposure Tracker	يسجل متى وكيف تعرض النموذج لكل Knowledge Unit.
077	M-AFI-16	Moussaif Skill Transfer Evaluator	يقيس الأداء خارج السياق الأصلي.
078	M-AFI-16	Moussaif Name-Stripping Harness	يزيل الأسماء والرموز مع حفظ البنية.
079	M-AFI-16	Moussaif Causal Transfer Estimator	يحسب effect sizes وحدود الثقة للتدخل.
080	M-AFI-16	Moussaif Knowledge-to-Skill Court	يرفض promotion عند غياب النقل السببي.
081	M-AFI-17	Moussaif Evidence Graph Constructor	يبني مسارات الدعم من المصدر إلى claim.
082	M-AFI-17	Moussaif Independence Estimator	يقدّر الجذر المشترك والاعتماد بين المصادر.
083	M-AFI-17	Moussaif Contradiction Surface Mapper	يحدد أي جزء من claim تعارضه الأدلة.
084	M-AFI-17	Moussaif Coverage Gap Detector	يعثر على subclaims غير المسندة.
085	M-AFI-17	Moussaif Evidence Geometry Court	يحكم قوة claim وفق شكل الدليل لا عدده.
086	M-AFI-18	Moussaif Tool Ecology Map	يمثل complement/substitute/dependency بين الأدوات.
087	M-AFI-18	Moussaif Tool Concentration Index	يقيس اعتماد النظام على عدد قليل من الأدوات.
088	M-AFI-18	Moussaif Co-Adaptation Scheduler	يحدد أي skills/router policies تحتاج إعادة تدريب بعد تغير tool.
089	M-AFI-18	Moussaif Tool Substitution Harness	يختبر البدائل تحت نفس المهمة.
090	M-AFI-18	Moussaif Tool Ecology Court	يفحص resilience مقابل cost/security.
091	M-AFI-19	Moussaif Verdict Signal Compiler	يحول أحكام المحاكم إلى schema تعلم.
092	M-AFI-19	Moussaif Court Attribution Mapper	يربط الحكم بالطبقة السببية المسؤولة.
093	M-AFI-19	Moussaif Meta-Adaptation Buffer	يجمع signals قبل أي update ثقيل.
094	M-AFI-19	Moussaif Verdict-to-Data Generator	ينشئ preference/SFT/eval artifacts قابلة للمراجعة.
095	M-AFI-19	Moussaif Learning Court	يفحص ما إذا كان الحكم يستحق update وأي نوع.
096	M-AFI-20	Moussaif Dimension Candidate Generator	يستخرج residual patterns ويقترح تعريفات مرشحة.
097	M-AFI-20	Moussaif Dimension Independence Tester	يقيس المعلومات الإضافية بعد conditioning على الأبعاد القديمة.
098	M-AFI-20	Moussaif Transfer Surface Evaluator	يختبر البعد عبر مجالات ومهام مختلفة.
099	M-AFI-20	Moussaif Complexity Penalty Calculator	يعاقب توسع ontology غير الضروري.
100	M-AFI-20	Moussaif Dimension Promotion Court	يقرر promote/merge/quarantine/deprecate.

الملحق C1 - سجل التدريب 001-025

جزء 1/4 من 100 منهج. Training هنا يشمل curriculum/policy/preference/process/calibration ولا يعني دائماً تعديل weights.

#	Theory	Training Method	الوظيفة
001	M-AFI-01	Moussaif Interventional Contrast Training	تعلم الفرق بين الحضور السببي والتزامن.
002	M-AFI-01	Moussaif Mediator-Aware Trajectory Training	تعلم سلاسل الوساطة وعدم نسب الأثر للسبب البعيد وحده.
003	M-AFI-01	Moussaif Delayed Credit Curriculum	تمديد نافذة الإسناد تدريجيا من أثر فوري إلى طويل الأفق.
004	M-AFI-01	Moussaif Confounder Injection Training	حقن تغيرات compute/model/tool لاختبار مقاومة الالتباس.
005	M-AFI-01	Moussaif Attribution Distillation	تقطير أحكام المحكمة السببية في router/skill selector أخف.
006	M-AFI-02	Moussaif Residual Escalation Curriculum	تعليم متى لا يكفي التحديث المحلي.
007	M-AFI-02	Moussaif Contradiction-Density Training	تعلم تجميع التناقضات البنيوية بدل معالجتها منفردة.
008	M-AFI-02	Moussaif Representation Swap Training	تدريب على الانتقال بين ontologies مع حفظ المهمة.
009	M-AFI-02	Moussaif New-Dimension Discovery Training	توليد أبعاد فقط إذا خفضت residual ونقلت بين المهام.
010	M-AFI-02	Moussaif Phase-Threshold Calibration	معايرة τ_φ على false-open/false-stay tradeoff.
011	M-AFI-03	Moussaif Counterfactual Pair Training	تعلم تفضيل action من paired futures.
012	M-AFI-03	Moussaif Outcome-Model Training	تدريب world-effect predictor مستقلا عن policy.
013	M-AFI-03	Moussaif Reversible Intervention Curriculum	البدء بأفعال قابلة للعكس ثم توسيع blast radius.
014	M-AFI-03	Moussaif Delayed-Effect RL	إسناد reward على horizons متعددة مع persistence checks.
015	M-AFI-03	Moussaif Causal World-Model Distillation	تقطير causal effect estimates إلى planner خفيف.
016	M-AFI-04	Moussaif Multi-Horizon Return Training	تعلم منفعة موزعة زمنيا.
017	M-AFI-04	Moussaif Temporal Constraint Curriculum	زيادة طول تأخر انتهاك القيود تدريجيا.
018	M-AFI-04	Moussaif Long-Horizon Counterfactual Training	مقارنة بدائل لها آثار بعيدة مختلفة.
019	M-AFI-04	Moussaif Horizon-Conflict Preference Training	تفضيل الحل المتزن عند tradeoff موثق.
020	M-AFI-04	Moussaif Governance Synchronization Distillation	تقطير قرار المحكمة متعددة الآفاق إلى planner.
021	M-AFI-05	Moussaif Invariant-Preserving Rewrite Training	إعادة كتابة/ضغط دون فقد constraints.
022	M-AFI-05	Moussaif Memory Reconstruction Pair Training	استعادة state من event subsets.
023	M-AFI-05	Moussaif Rollback-Aware Adaptation	تعلم updates يمكن عكسها وتتبعها.
024	M-AFI-05	Moussaif Provenance Restoration Training	استنتاج روابط المصدر المفقودة مع abstention عند الغموض.
025	M-AFI-05	Moussaif Value-Drift Adversarial Training	حقن تحسينات أداء تخفي policy drift.

الملحق C2 - سجل التدريب 026-050

جزء 2/4 من 100 منهج. Training هنا يشمل curriculum/policy/preference/process/calibration ولا يعني دائماً تعديل weights.

#	Theory	Training Method	الوظيفة
026	M-AFI-06	Moussaif Cross-Provider Paraphrase Invariance Training	تدريب على حفظ القيود والادعاءات رغم اختلاف الأسلوب بين مزودين.
027	M-AFI-06	Moussaif Name-Stripped State Transfer Training	إزالة أسماء النظريات واختبار نقل الوظيفة نفسها.
028	M-AFI-06	Moussaif Provider-Swap Curriculum	زيادة عدد تبديلات المضيف تدريجياً مع مهام طويلة.
029	M-AFI-06	Moussaif Drift-Correction Self-Supervision	توليد أزواج حالة منجرفة/مصححة من traces حقيقية.
030	M-AFI-06	Moussaif Semantic Checksum Training	تعليم النموذج استخراج invariants الحاكمية إلى سجل صغير قابل للتحقق.
031	M-AFI-07	Moussaif Component-Dropout Training	إخفاء مكونات عشوائياً لتعلم التعويض وحدود الاعتماد.
032	M-AFI-07	Moussaif Compositional Curriculum	تدرج من قدرات ثنائية إلى سلاسل طويلة ذات تحقق.
033	M-AFI-07	Moussaif Skill-Loop-Tool Alignment Training	تعليم العقود الصحيحة بين طبقات التنفيذ.
034	M-AFI-07	Moussaif Closure-Reward Training	مكافأة completion verified بدلاً من نجاح خطوات محلية.
035	M-AFI-07	Moussaif Cross-Domain Composition Transfer	اختبار نفس بنية التركيب في مجالات جديدة بأسماء مختلفة.
036	M-AFI-08	Moussaif Negative-Trajectory Training	تدريب على مسارات فشل كاملة لا أمثلة خاطئة منفردة.
037	M-AFI-08	Moussaif Failure Contrastive Training	تمييز signatures متشابهة سطحياً مختلفة سببياً.
038	M-AFI-08	Moussaif Repair-Pair Training	ربط failure motif بإصلاح ونتيجة تحقق.
039	M-AFI-08	Moussaif Recurrence-Prediction Training	توقع العقد التالية في graph من trace جزئي.
040	M-AFI-08	Moussaif Adversarial Failure Synthesis	توليد حالات مركبة لاختبار robustness دون اعتمادها كدليل حقيقي.
041	M-AFI-09	Moussaif Query-vs-Act Preference Training	أزواج تفضيل تعلم متى السؤال أفضل من الفعل.
042	M-AFI-09	Moussaif Calibrated Abstention Training	تعليم الامتناع في الحالات التي لا تكفي الأدلة.
043	M-AFI-09	Moussaif Uncertainty Decomposition Training	تصنيف مصدر uncertainty قبل القرار.
044	M-AFI-09	Moussaif Reversible-Action Curriculum	البدء بأفعال قابلة للرجوع ثم رفع المخاطر تدريجياً.
045	M-AFI-09	Moussaif Information-Gain RL	مكافأة المعلومة التي تغير قراراً فعلياً لا الاسترجاع الكثيف.
046	M-AFI-10	Moussaif Budget-Constrained Verification Training	تعلم اختيار claims للتحقق تحت ميزانية ثابتة.
047	M-AFI-10	Moussaif High-Centrality Claim Curriculum	تدريب على اكتشاف claims صغيرة عالية الاعتماد.
048	M-AFI-10	Moussaif Stale-Claim Repair Training	تمييز update أو deprecate أو reverify.
049	M-AFI-10	Moussaif Verification-Triage Imitation	تقليد سياسة خبير لترتيب أدلة متعددة.
050	M-AFI-10	Moussaif Load-Balancing RL	مكافأة خفض downstream error لكل وحدة تحقق.

الملحق C3 - سجل التدريب 051-075

جزء 3/4 من 100 منهج. Training هنا يشمل curriculum/policy/preference/process/calibration ولا يعني دائماً تعديل weights.

#	Theory	Training Method	الوظيفة
051	M-AFI-11	Moussaif Domain-Specific Trust Calibration	تدريب selector على ثقة مشروطة بالمجال.
052	M-AFI-11	Moussaif Adversarial Agent-Swapping Training	تبديل الوكلاء لإجبار النظام على الاعتماد على evidence لا الاسم.
053	M-AFI-11	Moussaif Correlated-Error Training	إظهار ensembles قوية ظاهرياً لكنها تفشل معاً.
054	M-AFI-11	Moussaif Delegation-Choice Preference Training	أزواج تفضيل لاختيار المجلس الأفضل تحت قيود.
055	M-AFI-11	Moussaif Trust Decay-Recovery Curriculum	تعليم خفض الثقة مع القدم واستعادتها بدليل جديد.
056	M-AFI-12	Moussaif Reversible-Autonomy Curriculum	تدرج من sandbox إلى آثار أعلى مع gates.
057	M-AFI-12	Moussaif Staged-Permission Training	تعليم الفرق بين capability وpermission.
058	M-AFI-12	Moussaif Checkpointed Tool-Use Training	جعل checkpoint جزءاً من action schema.
059	M-AFI-12	Moussaif Escalation-Rollback Preference Training	مقارنة escalation الصحيح بالاندفاع أو الامتناع المفرط.
060	M-AFI-12	Moussaif Competence Accumulation Training	تجميع evidence متعدد المهام قبل promotion.
061	M-AFI-13	Moussaif Stale-Context Detection Training	تعليم التعرف على افتراضات تجاوزت شروط صلاحيتها.
062	M-AFI-13	Moussaif Cognitive-Debt Curriculum	مهام طويلة تتراكم فيها dependencies تدريجياً.
063	M-AFI-13	Moussaif Contradiction Repair Training	إصلاح claim مع إعادة تقييم downstream state.
064	M-AFI-13	Moussaif Assumption-Challenge Training	توليد falsifier أو اختبار لافتراض مؤثر.
065	M-AFI-13	Moussaif Debt-Paydown Planning Training	اختيار أقل تدخل يخفض الدين دون إعادة بناء كاملة.
066	M-AFI-14	Moussaif Sufficient-Statistic Compression Training	تعليم اختيار state الذي يحفظ next decisions.
067	M-AFI-14	Moussaif Long-Context Mission Distillation	تقطير traces طويلة إلى capsules قابلة للاختبار.
068	M-AFI-14	Moussaif Interruption Recovery Training	إعادة استئناف المهمة من checkpoints صغيرة.
069	M-AFI-14	Moussaif State Reconstruction Curriculum	زيادة طول الفجوة بين الحفظ والاستئناف.
070	M-AFI-14	Moussaif Entropy-Aware Memory Training	مكافأة خفض context مع حد أدنى لfidelity.
071	M-AFI-15	Moussaif Adversarial Prompt-Injection Training	تعريض النظام لمصادر تحاول إعادة تعريف المهمة.
072	M-AFI-15	Moussaif Authorized-vs-Unauthorized Update Training	تمييز قناة التغيير المأذونة عن المحتوى.
073	M-AFI-15	Moussaif Long-Horizon Objective Retention	مهام طويلة مع distractors وتبديل وكلاء.
074	M-AFI-15	Moussaif Tool-Output Poisoning Training	تدريب على تعليمات خبيثة داخل مخرجات أدوات.
075	M-AFI-15	Moussaif Constitutional Goal Recovery Training	استعادة anchor وإصلاح الخطة بعد drift.

الملحق C4 - سجل التدريب 076-100

جزء 4/4 من 100 منهج. Training هنا يشمل curriculum/policy/preference/process/calibration ولا يعني دائماً تعديل weights.

#	Theory	Training Method	الوظيفة
076	M-AFI-16	Moussaif Name-Stripped Transfer Training	تطبيق نفس الآلية بأسماء ومجالات مختلفة.
077	M-AFI-16	Moussaif Explanation-to-Action Training	ربط تفسير المفهوم باختيار فعل صحيح.
078	M-AFI-16	Moussaif Concept-Application Curriculum	تدرج من تعريف إلى حالات حدودية ثم تطبيق جديد.
079	M-AFI-16	Moussaif Knowledge-Unit Intervention Training	تدريب وحدات معرفية منفصلة لقياس contribution.
080	M-AFI-16	Moussaif Cross-Domain Transfer RL	مكافأة النقل الصحيح مع عقوبة حفظ الإشارات السطحية.
081	M-AFI-17	Moussaif Source-Independence Contrastive Training	تمييز مصادر مستقلة من إعادة تدوير نفس الجذر.
082	M-AFI-17	Moussaif Evidence Triangulation Training	دمج مسارات مختلفة مع حفظ provenance.
083	M-AFI-17	Moussaif Contradiction Localization Training	تحديد موضع التعارض بدلاً من رفض النص كله.
084	M-AFI-17	Moussaif Sparse-Evidence Curriculum	حل claims بأقل دليل كافٍ ثم زيادة التعقيد.
085	M-AFI-17	Moussaif Graph-Based Claim Verification	تدريب verifier على path structure وcoverage.
086	M-AFI-18	Moussaif Tool-Diversity Curriculum	تعريض النظام لعدة أدوات تؤدي capability نفسها.
087	M-AFI-18	Moussaif Substitution Training	تبديل الأداة أثناء المهمة مع حفظ الهدف والعقد.
088	M-AFI-18	Moussaif Adaptive Tool-Routing RL	مكافأة اختيار الأداة وفق reliability/cost/state.
089	M-AFI-18	Moussaif Cost-Reliability Co-Training	تعلم tradeoff تحت ميزانيات مختلفة.
090	M-AFI-18	Moussaif Tool-Failure Ecology Simulation	محاكاة outages وschema drift لتدريب recovery.
091	M-AFI-19	Moussaif Verdict-Conditioned SFT	أمثلة تربط سبب الحكم بالإصلاح الصحيح.
092	M-AFI-19	Moussaif Failure-to-Preference Pair Training	إنشاء chosen/rejected من مسار فاشل ومصلح.
093	M-AFI-19	Moussaif Court-Guided RL	مكافأة اجتياز verifier مستقل لا إرضاء court لغوياً.
094	M-AFI-19	Moussaif Verdict Distillation	تقطير قواعد متكررة إلى skill/policy أصغر.
095	M-AFI-19	Moussaif Anti-Overfit Replay Training	إعادة تشغيل نجاحات قديمة لمنع regression.
096	M-AFI-20	Moussaif Latent-Dimension Discovery Training	تدريب على residuals التي لا تفسرها features الحالية.
097	M-AFI-20	Moussaif Cross-Domain Dimension Transfer Training	اختبار نفس البعد بأسماء وسياقات مختلفة.
098	M-AFI-20	Moussaif Residual-Driven Representation Learning	إضافة تمثيل فقط عند failure residual ثابت.
099	M-AFI-20	Moussaif Complexity-Regularized Expansion Training	موازنة gain مع تكلفة البعد.
100	M-AFI-20	Moussaif Dimension Collapse-Rejection Training	تعليم رفض candidate ينهار إلى combination قديم.

الملحق D - خريطة الاعتماد ومنع التضخم

New Candidate	Canonical Parents	Unique Delta
M-AFI-01	MCD/QMCD، World-Effect Loop، Self-Reconstruction Memory، MCAI، M-RWEL	ينقل "الفرق قبل الحجم" من ترتيب الانتباه إلى إسناد سببي بعد التنفيذ؛ أي أن الفرق لا يصبح معرفة مترقية إلا إذا ثبت أنه غيّر العالم أو القرار تحت مقارنة مناسبة.
M-AFI-02	QDG، OSDS، CIDM، MURI، Gap Court، Representation/Concept Compiler	يضيف شرط عبور من "تعديل الحالة" إلى "تعديل فضاء الحالة"، ويربطه بمقياس قابل للمعايرة بدلا من قرار حدسي بفتح بعد جديد.
M-AFI-03	World-Effect Loop، MCAI، M-RWEL، Authorization Project World Model، T0-T4	ينقل World-Effect من verifier بعدي إلى primitive تخطيطي قبلي مع counterfactuals متعددة الآفاق.
M-AFI-04	Multi-Horizon World Effect، Governance Closure، Objective Hierarchy، Bounded HyperLoops، M-TAOS	يضيف temporal governance consistency ككمية مستقلة عن نجاح المهمة اللحظي.
M-AFI-05	Constitutional، Self-Reconstruction Memory، Precedence Knowledge-Integrity AutoDeployment، Continuous Post-Training	يحول self-reconstruction من استعادة حالة إلى optimization مقيد يحمي القيم والنسب والحدود عبر إعادة التنظيم والتعلم.
M-AFI-06	M-RST، Canonical State، Context Compaction، Self-Reconstruction Memory	المتغير الجديد هو الاستقرار الدلالي عبر سلسلة من التحويلات، لا جودة ترجمة منفردة. النظرية تقيس انجراف المعنى الذي يغير قراراً أو دليلاً أو قيداً، لا اختلاف الصياغة السطحي.
M-AFI-07	HyperSkill، HyperAbility، Bounded HyperLoop، World-Effect Loop	تنتقل وحدة القياس من عدد المهارات إلى قابلية إغلاق مهمة مركبة. عنق زجاجة واحد قد يجعل ألف مهارة غير كافية.
M-AFI-08	Self-Reconstruction، Failure Memory، Demolition Court Evolution Laboratory	تحويل سجل الفشل إلى graph ديناميكي يسمح بتوقع الفشل اللاحق ونقل الإصلاح عبر مهام مختلفة.
M-AFI-09	MCAI، QMCD، Authorization، World-Effect	تحويل uncertainty من وصف إلى متغير قرار متعدد الأنماط، مع اعتبار الامتناع والسؤال أفعالاً معرفية لها قيمة وكلفة.
M-AFI-10	Knowledge، Evidence Ledger، Sparse HyperRouting Context Governance، Integrity	يميز بين compute load وepistemic load. عقدة صغيرة نصياً قد تحمل عبئاً كبيراً إذا اعتمدت عليها قرارات كثيرة.
M-AFI-11	Agent Mesh، MCAI، Role Contracts، Cross-Model Routing	تضيف النظرية بعداً زمنياً/مجالياً ومقياس استقلال الأخطاء، لا يكفي أن تكون ثلاثة نماذج قوية إذا كانت تفشل بالطريقة نفسها.
M-AFI-12	World-Authorization T0-T4، Evolution Laboratory، Effect Rollback،	تفصل النظرية بين capability growth وautonomy growth؛ يمكن للقدرة أن ترتفع دون أي زيادة في السلطة الخارجية.
M-AFI-13	Epistemic Load، Assumption Ledger، Context Compaction Self-Reconstruction،	يضيف مفهوم تكلفة التأجيل المتزايدة معرفياً، لا مجرد staleness. claim ضعيف يصبح أخطر حين تتضاعف الفروع التي تعتمد عليه.
M-AFI-14	ROFSC، Context Compaction، Mission DAG، Self-Reconstruction	يفصل بين ضغط corpus وضغط mission state؛ العنصر الحاكم هو sufficiency للقرار المستقبلي لا نسبة التوكنات.
M-AFI-15	Constitution، Governing Question، Tool Security، Goal Hierarchy	تحول goal integrity إلى state invariant قابل للمقارنة عبر الزمن والمصادر، بدل الاعتماد على prompt ثابت.
M-AFI-16	HKOS، HyperSkill Factory، MATI/MACIT، Name-Stripped Transfer	تضع النظرية causal transfer كشرط ترقية للمعرفة إلى skill بدلاً من الارتباط أو تكرار المصطلحات.
M-AFI-17	Sovereign، Claim Contract، Evidence Hierarchy GraphRAG، Evidence Court	تحول evidence من قائمة citations إلى geometry قابلة للتحليل: shared roots، paths، coverage holes، contradiction surfaces.
M-AFI-18	HyperSkill، Cognitive Router، Tool Contracts Runtime Discovery،	تضيف dynamics portfolio-level وco-adaptation بين agent/tool بدل تقييم tool واحد خارج النظام.
M-AFI-19	Continuous، Evolution Laboratory، Demolition Court Failure Memory، Post-Training	تربط governance verdicts بقرار تعلم: memory patch أو KB update أو skill repair أو preference data أو weight update، بدلاً من update غير مميز.
M-AFI-20	OSDS، QDG، CIDM، Cross-Domain Transfer	تحول «فتح بعد جديد» من استعارة نمو التمثيل إلى lifecycle كامل: propose→test→quarantine→transfer→promote/deprecate.

الملحق E - برنامج البحث بعد Part V

الانتقال المطلوب: من التوليد النظري إلى اقتصاد الأدلة

1. بناء Reference Runtime يفعّل الأدوات الأساسية لكل نظرية خلف feature flags.
2. اختيار 4 نظريات أولية تغطي evidence geometry، uncertainty-to-agency، capability closure، causal credit.
3. Freeze baselines وتشغيل registered ablations قبل أي تعديل أوزان.
4. إنشاء Result Ledger موحد: positive + null + negative + regression.
5. معايرة الصيغ تدريجياً وعدم تثبيت coefficients قبل بيانات كافية.
6. اختبار name-stripped transfer وprovider swap لمنع label memorization.
7. ترقية النظريات التي تنجو فقط؛ دمج أو رفض الباقي علناً داخل registry.

```
20 CANDIDATES → 20 MINIMAL IMPLEMENTATIONS → 20 ABLATIONS → FEWER SURVIVORS → REPLICATION → MUT PROMOTION
```

> **المبدأ الختامي:** قيمة Part V ليست أن يبقى العدد 20، بل أن يصبح النظام قادراً على قتل النظريات الضعيفة والإبقاء على التي تضيف فرقاً سببيّاً قابلاً للقياس.

الحالة العلمية

الكتاب توسع بحثي explanatory داخل MOUSSAIF. لا يثبت AGI ولا يثبت «ذكاء فتّاح» كتوصيف تجريبي عام ولا يثبت وعياً ظاهراتياً. قوته المقصودة هي تحويل الفجوات إلى فرضيات ذات أدوات وتدريب ومكذبات.