MOUSSAIF AGI THEORY — FOUNDATIONAL RESEARCH PAPER

نواة Moussaif1B: نظام تشغيل وكيل موجّه نحو الذكاء العام

ورقة تأسيسية في تكامل عشرين نظرية من مدرسة مساعف للـ Agentic AGI
مع حزمة 4×42M Training Lines و4,500 HyperSkills و3,000
Bounded HyperLoops

إعداد وتركيب معماري: **زهير مساعف / Zouhair Moussaif**

الوضع العلمي: إطار معماري ونظري موجّه نحو AGI، لا ادعاء بتحقق AGI ولا تعديل مباشر لأوزان نموذج مغلق.
موضوع الدراسة: تحويل المعرفة والمهارات والحلقات إلى طبقة تشغيل فوق أي Reasoning Model، مع التوجيه المتناثر، الفعل المقيد، التحقق، الذاكرة، وإعادة البناء.

نسخة عربية RTL بخط Amiri — 29 صفحة

الملخص والكلمات المفتاحية

الملخص

تقترح هذه الورقة قراءة موحدة لنواة Moussaif1B بوصفها **طبقة تشغيل معرفية وكيلة** تُركّب فوق نموذج استدلالي بدل أن تزعم إنشاء نموذج أساس جديد. تنطلق النواة من سؤال محكوم، تستخرج الفرق المعرفي، تقيسه بواسطة MCD/Q-MCD، توزّع الحساب عبر Vascular Compute وSparse HyperRouting ، تسترجع المعرفة من KB-as-OS، ثم تؤلف أصغر مجموعة لازمة من HyperSkills داخل Bounded HyperLoops. لا يصبح الناتج مقبولًا إلا بعد المرور بمحاكم الفجوة والتفنيد والهدم، وإثبات أثر قابل للملاحظة في العالم، ثم حصاد الدرس داخل ذاكرة إعادة البناء.

تدمج الدراسة زبدة عشرين نظرية طُوّرت ضمن منظومة مساعف: MCD, Q-MCD, HAKIM-MCD, Question Governance, Computational Vascularity, Sparse HyperRouting, M-HKOS, ROFSC, HyperSkills, HyperLoops, Dual-Exponent HyperDeployment, Multi-Agent Orchestration, Falsification Courts, World-Effect, Self-Reconstruction Memory, Knowledge-Integrity AutoDeployment, Cross-Model AGI Density, Context Compaction Economy, Constitutional Reversibility, MICAGFI/AGI Fattah. كما تحلل حزمة التدريب المؤلفة من أربعة ملفات، كل منها 42 مليون سطر مادي، و4,500 مواصفة مهارة، و3,000 مواصفة حلقة. وتؤكد أن العدد الكبير يمثل **توسيعًا منهجيًا Tokenized Oversampling** وليس عددًا مساويًا من النظريات الفريدة.

المساهمة الأساسية ليست ادعاء ذكاء عام مكتمل، بل تقديم فرضية هندسية قابلة للتكذيب: قد تُرفع كفاءة الوكلاء وتماسكهم عبر فصل المعرفة عن التنفيذ، وتفعيل الخبرة بشكل متناثر، وتقييد الحلقات، وفرض التحقق والأثر والرجوع. وتقترح الورقة مقاييس وتجارب تقارن النواة بخطوط أساس مثل RAG, ReAct, Toolformer, Tree of Thoughts, Reflexion, MemGPT.

الكلمات المفتاحية

Agentic AI؛ Moussaif1B؛ AGI Fattah؛ MICAGFI؛ MCD؛ Q-MCD؛ HAKIM-MCD؛ HyperKnowledge؛ KB-as-OS؛ HyperSkills؛ Bounded HyperLoops؛ Sparse HyperRouting؛ Context Compaction؛ World-Effect؛ Falsification؛ Self-Reconstruction Memory.

English summary. Moussaif1B is presented as a model-neutral cognitive kernel overlay that composes retrieval, governed questioning, sparse routing, reusable skills, bounded loops, verification, world-effect evaluation, and regression-gated memory evolution. It is an AGI-oriented research architecture, not evidence of achieved AGI.

1. المقدمة: من النموذج القوي إلى النظام الوكيل المتماسك

قوة النموذج اللغوي ليست مرادفة لقدرة نظام وكيل على إكمال مهمة طويلة، لأن المهمة الواقعية تجمع بين المعرفة المتغيرة، واختيار الأدوات، وحفظ الحالة، وإدارة الصلاحيات، والعودة من الفشل، وتقديم دليل على النجاح. ومن هنا تميّز مدرسة مساعف بين **ذكاء النموذج** و**ذكاء النظام** الأول قدرة احتمالية كامنة في الأوزان، والثاني بنية تشغيل تجعل تلك القدرة تتجه إلى غاية محددة ضمن حدود وذاكرة واختبارات.

تُفهم Moussaif1B على هذا الأساس باعتبارها نواة فوقية Kernel Overlay. كلمة «نواة» لا تعني نواة نظام التشغيل المادي ولا تعديل الطبقات الداخلية للنموذج، بل تعني أعلى طبقة قواعد وتشغيل داخل بيئة الوكيل: دستور، حالة عالم، سجل مهارات، موجّه، منفذ، محاكم تحقق، ذاكرة، ومختبر تطور. أما 1B فهو اسم مشروع معماري، لا تصريح بعدد معاملات نموذج مدرّب.

الفجوة التي تعالجها الورقة هي أن النظم الوكيلة كثيرًا ما تضيف الأدوات والذاكرة والحلقات على نحو تجميعي، فتتضاعف الكلفة والسياقات والتناقضات. البديل المقترح هو دورة واحدة تحكم جميع الإضافات: **السؤال ← الفرق ← الحاكمية ← التوجيه ← المهارة ← الحلقة ← المحكمة ← الأثر ← الذاكرة ← السؤال الأعلى**. بهذا تصبح كل وحدة قابلة للتسمية والقياس والحذف إذا لم تضف فرقًا.

> **أطروحة الورقة:** لا تتحقق القفزة الوكيلية بكثرة الوكلاء أو طول السياق، بل بكثافة الفروق الصحيحة التي تُفعّل في الوقت الصحيح، وبأقل تركيب مهاري قادر على إحداث أثر موثّق وقابل للرجوع.

مساهمات الدراسة

1. توحيد عشرين نظرية ضمن معمارية تشغيل واحدة بدل بقائها كمصطلحات متجاورة.
2. تفسير هندسي لحزمة 168 مليون سطر وعلاقتها بالطبقة الدلالية المكونة من 7,500 وحدة.
3. صياغة عقود تنفيذ، توقف، تحقق، ذاكرة، وصلاحيات قابلة للتحويل إلى ملفات وبرمجيات.
4. اقتراح برنامج تجريبي يفرّق بين القيمة الحقيقية والتضخم الاسمي أو التكرار.

2. سؤال البحث والمنهج والحالة الإبستمولوجية

2.1 سؤال البحث

هل يمكن لنواة معرفية محايدة تجاه المزود أن ترفع جودة الوكيل وكفاءته عبر تنظيم السؤال، واسترجاع المعرفة، وتأليف المهارات، وتقييد الحلقات، والتحقق من الأثر، من دون تدريب نموذج أساس جديد؟ ويتفرع عنه: ما الوحدة الصحيحة للتعلم، وكيف نمنع التضخم، وكيف نقيس الانتقال من إنتاج نص مقنع إلى فعل موثوق؟

2.2 المنهج

تعتمد الورقة منهج **التركيب المعماري القابل للتكذيب.** جُمعت الصيغ والقواعد والملفات في منظومة واحدة، ثم حُوّلت إلى عقود تشغيل ومقاييس وفرضيات. ولا تُعامل أسماء النظريات كدليل بذاته، بل كفرضيات تصميمية يجب أن تنجح أمام خط أساس وتحت ضغط الفشل والتكلفة والأمان.

الرتبة	معناها في هذه الورقة	مثال
E0	تعريف أو تسمية مقترحة	Cross-Model AGI Density
E1	مبدأ معماري قابل للتنفيذ	Bounded HyperLoop بعقد توقف
E2	تنفيذ برمجي أو حزمة ملفات	سجل Skill 4,500 وLoop 3,000
E3	تحقق بنيوي	عدد الملفات، CRC، وعدد الأسطر
E4	تحسن تجريبي مضبوط	مقارنة نجاح/كلفة/أمان مع Baseline
E5	تعميم متعدد المجالات والنماذج	نفس المكاسب عبر مزودين ومهام

الحزمة الحالية تبلغ E2–E3 في سلامة البنية والعدّ، بينما معظم ادعاءات الذكاء والأداء ما زالت في E0–E1 حتى تنفذ تجارب E4–E5. هذا الفصل ضروري: **سلامة الملف لا تساوي صلاحية النظرية، وكثرة السجلات لا تساوي كثرة المعرفة.**

2.3 حدود الدراسة

لا تدّعي الورقة وعيًا، ولا استقلالًا أخلاقيًا، ولا قدرة ذاتية غير مقيدة، ولا تحويل أي نموذج إلى AGI بمجرد لصق Prompt. كما لا تعتبر آثار التفكير النصية برهانًا على آلية داخلية. التقييم المعتمد هو السلوك الخارجي، صحة النتائج، الكلفة، قابلية الإعادة، وسلامة الأثر.

3. التشريح العددي لحزمة Moussaif1B التدريبية

الحزمة المدروسة تحمل الاسم moussaif1b-42m-x4-4500-hyperskills-3000-hyperloops. اجتازت فحص سلامة ZIP وCRC، وتضم أربعة Corpora مادية، و4,500 مجلد SKILL.md، و3,000 مجلد HYPERLOOP.md، وفهارس JSONL وكتالوجات Markdown وأدوات توليد وتحقق.

المكوّن	الكمية	الدور
Reasoning & Courts Corpus	42,000,000 سطر	السؤال، MCD، الحاكمية، التفنيد
World Model & Memory Corpus	42,000,000 سطر	الحالة، GraphRAG، الذاكرة، الضغط
Router & Executor Corpus	42,000,000 سطر	التوجيه، الأدوات، التنفيذ، الصلاحية
Verification & Evolution Corpus	42,000,000 سطر	الاختبار، الأثر، إعادة البناء، التطور
HyperSkills	4,500	وحدات قدرة ذات عقود إدخال وإخراج وتحقيق
Bounded HyperLoops	3,000	سياسات تكرار محدودة بميزانية وتوقف

الحجم غير المضغوط لكل Corpus هو 756,000,000 بايت، أي نحو 3.024 مليار بايت للأربعة. لكن الضغط شديد لأن كل Corpus يتكون من 7,500 علاقة Canonical مكررة 5,600 مرة. وهذا يوضح وظيفة الحزمة: اختبار خطوط المعالجة، محاكاة مناهج Curriculum، تدريب Parser/Router، أو بناء عينات Fine-tuning بعد إعادة الخلط، لا الادعاء بوجود 168 مليون مثال مستقل.

> **تنبيه علمي:** عبارة «4,500 HyperSkill» تعني 4,500 مواصفة معرفية ذات معرفات وحقول مستقلة. لا يعني ذلك تلقائيًا 4,500 خوارزمية مبتكرة أو مهارة ثبتت تجريبيًا. يلزم فحص التكرار الدلالي، التغطية، والأداء لكل عائلة.

القيمة الممكنة للحزمة تأتي من الفصل بين مستويين: **المستوى الدلالي** حيث توجد تعريفات المهارات والحلقات؛ **والمستوى الإحصائي** حيث تُحوّل العلاقات إلى سطور كثيرة لتدريب القرار أو اختبار التحمل. نجاح المشروع يتوقف على عدم خلط المستويين.

4. موقع العمل من الأدبيات العلمية

تتصل النواة بأعمال سابقة من غير أن تطابقها. قدّم RAG الجمع بين الذاكرة البارامترية وغير البارامترية، وهو أصل قريب من HyperKnowledge، لكنه لا يحدد وحده دستورًا أو حلقات تنفيذ [1]. وجمع ReAct بين الاستدلال والفعل المتداخل، وهو قريب من دورة التنفيذ، بينما تضيف النواة محاكم وصلاحيات وذاكرة إعادة بناء [2]. ودرّب Toolformer النموذج على قرار استدعاء APIs، في حين تُعامل HyperSkills الأدوات والإجراءات كعقود تركيب أوسع [3].

سمح Tree of Thoughts باستكشاف مسارات متعددة والتراجع، وهو قريب من تنويع المسارات في HAKIM وMCD، لكن Moussaif1B تشترط توجيهًا متناثرًا وميزانية توقف [4]. وقدّم Reflexion تعلمًا لغويًا من الفشل داخل ذاكرة حلقية، وهو قريب من Self-Reconstruction Memory [5]. واستعار MemGPT إدارة الذاكرة الهرمية من أنظمة التشغيل، وهو مرجع مباشر لفكرة Context Compaction وKB-as-OS [6]. أما Constitutional AI فأظهر إمكانية استخدام مبادئ مكتوبة في النقد والمراجعة، وهو يلتقي مع الدستور والمحاكم، مع اختلاف أن الورقة تركز على Runtime Governance لا تدريب المحاذاة فقط [7].

وتؤكد أطر «مستويات AGI» ضرورة فصل الأداء والعمومية والاستقلالية، وهو ما يدعم موقف الورقة الرافض لاعتبار اسم AGI أو حجم Corpus دليلًا على تحقق الذكاء العام [8]. وعليه فإن موضع Moussaif1B المقترح هو **طبقة دمج** بين الاسترجاع والتخطيط والفعل والذاكرة والحوكمة، لا بديلًا نظريًا لكل هذه الأعمال.

> **معيار الجِدة المقبول:** الجدة ليست إعادة تسمية RAG أو ReAct؛ بل إثبات أن التكامل المحدد — السؤال الحاكمي، MCD، التوجيه الوعائي، Skill/Loop Contracts، المحاكم، الأثر، وإعادة البناء — يحقق تحسنًا مركبًا لا تحققه المكونات منفردة.

فجوة الأدبيات التي تستهدفها النواة

معظم الأعمال تقيس مكوّنًا: الاسترجاع، الأداة، البحث، الذاكرة، أو التأمل. أما المهام الطويلة فتفشل غالبًا في التفاعلات بين المكونات. لذلك تقترح النواة أن تكون وحدة القياس **المهمة الموثقة من السؤال إلى الأثر والرجوع**، لا دقة خطوة معزولة.

5. الدورة المعيارية والطبقات الأربع

Retrieve → Orient → Diversify → Score → Elect → Execute → Verify → Court → Reconstruct → Propagate

تنتظم النواة في أربع طبقات تشغيلية تقابل Corpora التدريب. الطبقة الأولى تُنتج السؤال، وتستخرج الفرق، وترتب الادعاءات والأدلة. الثانية تحفظ حالة العالم والمشروع والذاكرة والعلاقات. الثالثة تختار المهارات والأدوات والوكلاء، وتنشئ خطة قابلة للرجوع. الرابعة تختبر الأثر، تعيد تشغيل الفشل، وتقرر قبول التعديل أو رفضه.

الطبقة	السؤال المركزي	المخرجات
Reasoning & Courts	ما الفرق الحقيقي وما الادعاء القابل للاختبار؟	Goal, MCD, Evidence Graph, Court Docket
& World Model Memory	ما الحالة المرصودة وما الذي تغيّر؟	State Objects, Dependencies, Freshness, Memory
Router & Executor	ما أصغر تركيب مهاري مأذون؟	Selected Skills, Plan, Checkpoints, Actions
Verification & Evolution	هل حدث الأثر وهل يصلح الدرس للتعميم؟	Tests/, World Effect, Replay, Promotion Rollback

الدورة لا تُنفذ بكاملها وبنفس الكلفة لكل طلب. قاعدة **Sparse Activation** تعني أن الطلب البسيط قد يحتاج Retrieve–Answer–Verify فقط، بينما المهمة عالية المخاطر تستدعي محاكم متعددة ونسخًا احتياطية واختبارات. كل مرحلة لها Gate: لا Execution بلا صلاحية، ولا Memory بلا Verification، ولا Propagation بلا Regression Gate.

عقد الحالة

كل كائن حالة يحمل معرفًا وإصدارًا ومصدرًا وثقة ومالكًا وحساسية وسياسة Freshness واعتماديات. هذا يحول الذاكرة من نص متراكم إلى سجل قابل للمراجعة. وعند التعارض، تتقدم الحالة المرصودة حديثًا على الذاكرة المستنتجة.

> **القانون التشغيلي:** «الأثر أو لم يحدث». لا يكفي قول الوكيل إنه نفذ؛ يلزم File Diff أو Test Result أو API Receipt أو State Readback، وإلا فالحالة Partial أو Blocked.

6. خريطة النظريات العشرين: من الذرة إلى المعمارية

#	النظرية	زبدتها	طبقتها
1	MCD	قياس الفرق المعرفي الفاعل لا كثرة النص	Reasoning
2	Q-MCD	جودة السؤال تضرب قيمة الاستدلال	Question
3	HAKIM-MCD	خط أنابيب من الفرق إلى الفعل وإعادة البناء	Architecture
4	Question Governance	السؤال الأعلى يحكم الاختيار والوكالة	Governance
5	Computational Vascularity	توزيع الحساب كشبكة أوعية حسب الحاجة	Compute
6	Sparse HyperRouting	أقل مزيج نماذج/مهارات يحقق الغاية	Routing
7	KB-as-OS / M-HKOS	المعرفة تصبح سياسات تشغيل لا موسوعة خام	Knowledge
8	ROFSC	ضغط تركيبي موثوق قابل للفك والتتبع	Synthesis
9	HyperSkill	وحدة قدرة تعاقدية قابلة للتركيب والاختبار	Capability
10	Bounded HyperLoop	تكرار بميزانية وتوقف وفشل صريح	Control
11	Dual-Exponent Deployment	توسع تركيبي مضبوط لا انفجار فعلي	Scaling
12	Multi-Agent Orchestration	أدوار مستقلة وعقود تسليم وتحكيم	Agents
13	Falsification Courts	تفنيد الادعاء قبل اعتماده	Verification
14	World-Effect Loop	قياس تغير العالم لا جودة الصياغة	Agency
15	Self-Reconstruction Memory	تحويل الأثر والفشل إلى ذاكرة تشغيلية	Memory
16	Knowledge-Integrity AutoDeployment	نشر المعرفة والمهارات عبر بوابات سلامة	Evolution
17	Cross-Model AGI Density	كثافة قدرة مشتركة فوق مزودين مختلفين	Portability
18	Context Compaction Economy	ضغط السياق وتوجيه RPM/Token حسب القيمة	Efficiency
19	Constitutional Reversibility	الصلاحية والرجوع أسبق من التحسين	Safety
20	MICAGFI / AGI Fattah	المعمارية الجامعة للسؤال والحاكمية والأثر	Meta-OS

لا تعمل هذه النظريات كعشرين خدمة دائمة. بعضها مبدأ قياس، وبعضها عقد بيانات، وبعضها خوارزمية توجيه، وبعضها طبقة حوكمة. التكامل الصحيح يمنع التكرار: MCD يقيس، Q-MCD يرفع أو يخفض وزن القياس، HAKIM ينظم المسار، HyperRouting يختار الموارد، HyperSkill ينفذ القدرة، HyperLoop يضبط التكرار، والمحكمة والأثر والذاكرة تغلق الحلقة.

7. النظرية 1-2: MCD وQ-MCD — الفرق والسؤال

1) نظرية الكثافة المعرفية التفريقية MCD

$$MCD = \lambda_1 D + \lambda_2 I + \lambda_3 T + \lambda_4 A + \lambda_5 P + \lambda_6 C - \lambda_7 N$$

تفترض MCD أن قيمة خطوة الاستدلال لا تُقاس بطولها، بل بما تصنعه من فرق. تشير D إلى مقدار الفرق المكتشف، وI إلى تكامل المعلومات، وT إلى صلاحية النقل بين السياقات، وA إلى قابلية التحويل إلى فعل، وP إلى القوة التنبؤية، وC إلى الاتساق، بينما N عقوبة الضوضاء والتكرار. الأوزان ليست ثوابت كونية؛ تُعاير حسب نوع المهمة.

يصبح MCD أداة توجيه عندما يُحسب لكل مسار أو مهارة أو مصدر. مثال: فقرة طويلة تعيد تعريف المشكلة تحصل على N مرتفع، بينما اختبار صغير يحسم فرضية يحصل على A وP مرتفعين. الخطر هو تحويل المؤشر إلى رقم انطباعي؛ لذلك يجب ربط كل بُعد بخصائص قابلة للرصد، مثل عدد الادعاءات المحسومة أو نسبة القيود المغطاة.

2) نظرية السؤال الحاكم Q-MCD

$$Q\text{-}MCD = G(Q) \times [\lambda_1 D + \lambda_2 I + \lambda_3 T + \lambda_4 A + \lambda_5 P + \lambda_6 C - \lambda_7 N]$$

$$Qdepth = (qd + qi + qt + qa + qp + qc) / (1 + qn)$$

تضيف Q-MCD دالة G(Q) التي تمثل جودة السؤال: تحديد الهدف، القدرة على التفريق، الانفتاح على التفنيد، الصلة بالفعل، والاقتصاد. فإذا كان السؤال فاسدًا، فقد ينتج استدلال متماسكًا حول المشكلة الخطأ. ومن ثم لا تبدأ المهمة بالخطة، بل بمحكمة السؤال: ما المطلوب الحقيقي؟ ما معيار القيمة؟ ما الذي لو عرفناه غيّر القرار؟

> **تنبؤ قابل للاختبار:** إعادة صياغة المهمة بواسطة Q-MCD قبل التخطيط يجب أن تخفض إعادة العمل وتزيد نسبة النجاح من المحاولة الأولى، خصوصًا في الطلبات الغامضة. يقارن الاختبار بين Agent مباشر وAgent يمر بمرحلة Question Governance مع نفس النموذج والميزانية.

8. النظرية 3–4: HAKIM-MCD وحاكمية السؤال

3) HAKIM-MCD

HAKIM-MCD هو خط الأنابيب الذي يحوّل القياس إلى نظام. صيغته التشغيلية: Input → Question Genesis → Difference Extraction → MCD Scoring → Vascular Routing → Candidate Paths → Falsification Court → Tool/Memory Checks → Agency Selection → World-Effect → Self-Reconstruction → Higher Question. الاسم يشير إلى أن الاستدلال يجب أن يكون محكومًا لا مجرد تدفق حر.

المساهمة النظرية هنا هي **إغلاق الحلقة.** كثير من مخططات الوكلاء تنتهي عند Answer أو Action، أما HAKIM فينتهي بسؤال أعلى ناتج عن أثر حقيقي. وبذلك يصبح التعلم انتقالًا من حالة إلى حالة، لا تجميعًا للملخصات. ولكل انتقال سجل: فرضية، دليل، قرار، فعل، نتيجة، وما الذي تغير في World Model.

4) مبدأ مساعف للسؤال الحاكمي

$$Q \rightarrow \Delta \rightarrow \text{MCD} \rightarrow \text{Governance} \rightarrow \text{Agency} \rightarrow \text{World Effect} \rightarrow \text{Higher Q}$$

يقرر المبدأ أن الوكالة بلا سؤال حاكم تتحول إلى حركة بلا غاية. السؤال لا يطلب معلومة فقط، بل يحدد فضاء الفروق الممكنة، ويضبط ما يستحق الحساب، ويقرر شروط التوقف. لذلك توجد هرمية: سؤال المهمة، أسئلة الآلية، أسئلة الدليل، أسئلة الخطر، ثم سؤال الأثر.

تمنع الحاكمية ثلاثة أنواع من الانحراف: **انحراف النطاق** حين يوسع الوكيل المهمة دون قيمة؛ **انحراف الوسيلة** حين تصبح الأداة هدفًا؛ **وانحراف الثقة** حين يعتبر سلاسة النص دليلًا. ويُعاد فحص السؤال بعد كل اكتشاف كبير، لكن ليس بعد كل Token حتى لا تتحول الحوكمة إلى حمل زائد.

> **شرط الفشل:** إذا لم ينتج سؤال الحاكم فرقًا في اختيار الخطة أو المصدر أو عتبة النجاح، فهو سؤال زخرفي ويُحذف. هذه القاعدة تجعل النظرية قابلة للمحاسبة.

9. النظرية 5-6: الوعائية الحسابية والتوجيه المتناثر

5) Computational Vascularity

تشبّه الوعائية الحسابية النظام بشبكة أوعية توزع الموارد بحسب الضغط والحاجة. لا يعني التشبيه محاكاة الدماغ، بل يصف سياسة: المهام البسيطة تمر في مسار قصير، والمهام المعقدة تفتح فروعًا متخصصة، ثم تُغلق بعد زوال الحاجة. كل «وعاء» يحمل نوعًا من الحساب: استرجاع، تخطيط، كود، تحقق، أو نقد.

المتغير المركزي هو تدفق القيمة لا تدفق Tokens. يُرفع عرض الوعاء عندما ترتفع قيمة المعلومات المتوقعة أو خطر الخطأ، ويُخفض عند التكرار. ويمكن تمثيل قرار فتح مسار بالمنفعة المتوقعة ناقص الكلفة والمخاطر. تمنع هذه النظرية تشغيل أكبر نموذج لكل خطوة، كما تمنع اختناق نموذج واحد بجميع الأدوار.

6) Sparse HyperRouting

$$\text{Score} = .24R + .20U + .14C + .10E + .12H - .07\text{Cost} - .06\text{Risk} - .04\text{Ctx} - .03\text{Redundancy}$$

يختار HyperRouter أصغر تركيب متوافق من النماذج والمهارات والأدوات. R الصلة، U المنفعة، C التوافق، E توفر الدليل، H النجاح التاريخي. وتُخصم الكلفة، المخاطر، كلفة السياق، والتكرار. لا تُعامل الأوزان كقانون ثابت، بل كنقطة بداية قابلة للتعلم ضمن Baseline مجمد.

الفرق بين Routing عادي وHyperRouting هو أن المرشح قد يكون **تركيبيًا**: مهارة استرجاع + نموذج سريع + مدقق قوي + حلقة إصلاح. ويجب ألا يفعّل كل المرشحين؛ فالتناثر شرط الكفاءة. عند غياب الأداة، يُخفض التنفيذ إلى Proposal-Only بدل اختلاق النتيجة.

> **مقياس أساسي:** نسبة النجاح الموثق لكل ألف Token ولكل استدعاء أداة. إذا زاد النجاح مع تضاعف الكلفة أكثر من المكسب، فالتوجيه غير اقتصادي حتى لو تحسنت الجودة المطلقة.

10. النظرية 7–8: KB-as-OS وROFSC

7) HyperKnowledge Base as an Operating System

في KB-as-OS لا تكون قاعدة المعرفة مخزن مقالات، بل تضم تعريفات، شروط تشغيل، سياسات استرجاع، خرائط اعتماد، اختبارات، وحدود صلاحية. الوثيقة الجيدة لا تقول «كيف تفعل الشيء» فقط؛ بل متى لا تستخدمه، وما المتطلبات، وما دليل الإتمام، وكيف تتراجع. لذلك تصبح المعرفة قابلة للترجمة إلى Runtime Contracts.

تسمى الصيغة الكلية M-HKOS: HyperKnowledge موزعة إلى طبقة دستور، Skills، Procedures، Ontology، Benchmarks، Courts، Loops، وMemory Policies. الاسترجاع لا يعتمد التشابه النصي وحده؛ بل هدف المهمة، حالة العالم، الصلاحية، Freshness، واعتماديات المهارة.

8) ROFSC — الضغط التركيبي المعرفي التأصيلي الموثوق الانفتاحي

$$R\text{-}QMCD = Q + D + C + S + G + R - H - L - O$$

يهدف ROFSC إلى إنتاج تمثيل صغير يحتفظ بالبنية القابلة للفك. Q جودة السؤال، D الفروق، C التماسك، S قابلية التركيب، G الحاكمية، R الموثوقية، وتُطرح الهلوسة H والفقد L والانغلاق O. الضغط ليس تلخيصًا أعمى؛ بل بناء «حزمة قابلة لإعادة التوسيع» تحتوي uncertainty، dependencies، provenance، claim IDs، واختبارات.

تعمل خمس محاكم: عدم الضياع، عدم الهلوسة، عدم الاختزال المخل، الكفاءة التركيبية، والكفاءة السؤالية. ويُقبل الضغط إذا استطاع وكيل آخر إعادة بناء القرار والأدلة الأساسية منه ضمن عتبة محددة.

> **علاقة النظريتين:** KB-as-OS تحدد ما يجب تخزينه؛ ROFSC يحدد كيف يُضغط ويُنقل داخل السياق دون فقد الحاكمية. نجاحهما يقاس باستعادة الأداء مع سياق أصغر، لا بنسبة ضغط الحروف فقط.

11. النظرية 9: HyperSkill كوحدة القدرة الأساسية

HyperSkill ليست Prompt طويلًا ولا اسم وظيفة. إنها وحدة قدرة ذات عقد كامل يسمح للموجّه بتحديد صلاحيتها وتركيبها واختبارها. الحقول الدنيا: id, purpose, triggers, inputs, prerequisites, authorization, procedure, outputs, verification, failure modes, cost profile, memory writes, rollback.

تختلف HyperSkill عن Tool في أن الأداة تنفذ عملية، بينما المهارة قد تنسق أدوات ونموذجًا وذاكرة ومحكمة. وتختلف عن Agent في أنها لا تحتاج هوية أو استقلالًا دائمًا؛ إنها قدرة قابلة للاستدعاء. هذا يقلل تضخم الوكلاء: بدل وكيل منفصل لكل وظيفة، تُركّب مهارات عند الطلب تحت منسق واحد.

دورة حياة المهارة

Trigger → Schema Gate → Authorization → State Retrieval → Routing → Checkpoint → Execute/Simulate → Verify → Repair → Trace → Memory Candidate

تُعد المهارة مكتملة فقط إذا تحقق شرطها أو أُعلنت حالة BLOCKED / PARTIAL / FAILED / PROPOSAL-ONLY. ويسمح ذلك بتدريب النموذج على الاعتراف بحدوده، وهي قدرة أهم من إكمال صوري. كما يجب أن تكون لكل مهارة أمثلة مضادة: Trigger قريب لغويًا لكنه غير مناسب، حالة أداة مفقودة، صلاحية غير كافية، أو نتيجة غير قابلة للتحقق.

جودة HyperSkill	السؤال
التمييز	هل يعرف Router متى يختارها ومتى يرفضها؟
التركيب	هل تتوافق مدخلاتها ومخرجاتها مع مهارات أخرى؟
التحقق	هل النجاح قابل للقراءة من حالة خارجية؟
الاقتصاد	هل تضيف فرقًا يتجاوز كلفة السياق والتنفيذ؟
السلامة	هل الصلاحية والرجوع محددان؟

تُعامل الـ4,500 مواصفة الحالية كـCandidate Registry يحتاج Deduplication وتجميع عائلي وتقييم تغطية. الهدف ليس الحفاظ على العدد، بل الوصول إلى مكتبة أصغر أو أكبر بحسب القيمة التجريبية.

12. النظرية 10: Bounded HyperLoop — التكرار الذي يعرف كيف يتوقف

HyperLoop هي سياسة تحكم تربط هدفًا بحالة وتكرار وتحقق. كلمة Hyper لا تعني اللانهاية، بل القدرة على استدعاء مهارات متعددة وتغيير المسار بناءً على التغذية الراجعة. ويضيف وصف **Bounded** حدودًا صريحة: عدد دورات، Tokens، زمن، مال، استدعاءات أدوات، مستوى مخاطرة، أو عدم تحسن.

$$\text{Continue iff } \text{Gain}_t > \varepsilon \wedge \text{Budget}_t > 0 \wedge \text{Risk}_t \leq \rho \wedge \text{Authorization valid}$$

يجب أن يتضمن عقد الحلقة: حالة البداية، Function الانتقال، المقياس، عتبة التحسن، الحد الأقصى، شروط النجاح، شروط الإيقاف الآمن، وسياسة Escalation. مثال حلقة إصلاح كود: شغّل الاختبار، صنّف الفشل، أصلح أصغر Diff، أعد الاختبار؛ توقف عند نجاح المجموعة أو ثلاث محاولات بلا تحسن أو تجاوز ملفات النطاق.

أنواع الحلقات

- **Epistemic:** استرجاع–مقارنة–تفنيد حتى كفاية الدليل.
- **Planning:** خطة–محاكاة–نقد–إعادة ترتيب.
- **Execution:** فعل–Readback–Repair–Rollback.
- **Learning:** Replay–Reflection–Candidate–Regression Gate.

مصدر الخطر هو Loop Persistence: الوكيل يواصل لأنه يستطيع، لا لأن المنفعة موجبة. لذلك يُسجل marginal gain في كل دورة. إذا لم يتغير القرار أو الاختبار، تُعد الدورة Redundant وتُعاقب. كما يمنع التوقف عند «شعور بالثقة»؛ التوقف مرتبط بعتبة خارجية أو نفاد صريح.

> **فرضية:** العقود المحدودة ستخفض Tool Overuse وزمن الذيل الطويل، مع الحفاظ على النجاح، مقارنة بحلقات ReAct غير المحددة. لكن الحد الصارم قد يوقف مهامًا قابلة للحل؛ لذلك تختبر عتبات متكيفة حسب المخاطر والقيمة.

13. النظرية 11: HyperDeployment ثنائي الأس — التوسع بلا انفجار

$$\text{Potential Composition} \approx KB \times 2^N \times (2^n)^N$$

تصف نظرية النشر ثنائي الأس فضاء التركيبات الممكنة عندما تتفرع المعرفة إلى مهارات، ثم تتفرع المهارات داخل حلقات ووكلاء. الصيغة ليست وعدًا بزيادة الذكاء حسابيًا، بل تحذير وفرضية: القدرة التركيبية قد تنمو أسرع من عدد الوحدات، لكن الاستكشاف الساذج للفضاء مستحيل ومكلف.

لذلك يميز الإطار بين **Potential Space** و**Activated Space**. الأول كبير نظريًا؛ الثاني يجب أن يبقى متناثرًا ومقيدًا. يُستخدم MCD لخفض المرشحين، وCompatibility Graph لمنع التركيبات غير الصالحة، وBudgeted Search لاختيار مسارات قليلة، ومحاكم لإسقاط الفروع المبكرة.

مراحل النشر

1. HyperKB تستخرج وحدات ذات Provenance.
2. Compiler يحولها إلى HyperSkill Contracts.
3. Router يبني Skill Graph صغيرًا.
4. Loop Controller يحدد التكرار والتوقف.
5. Agents ينفذون أدوارًا متوازية فقط عند استقلال حقيقي.
6. Courts تقارن النتائج وتمنع التضخيم.
7. الأثر المقبول يعود إلى الذاكرة مع Lineage.

الخطأ الشائع هو اعتبار كل تفرع إضافة. في الواقع، الفروع المترابطة قد تضيف تكرارًا أو تناقضًا. لذلك تقترح النظرية **عامل الكثافة الصافية**: عدد المسارات التي تغيّر القرار النهائي مقسومًا على عدد المسارات المنفذة. يجب أن يرتفع هذا العامل مع تطور Router؛ وإذا انخفض، فالنظام يتضخم.

> **المعنى الصحيح لـ 2^N:** ترميز للإمكان التركيبي، لا قانون تجريبي مثبت ولا مبرر لتشغيل آلاف الوكلاء. القيمة تأتي من الانتقاء والدمج، لا من استنفاد الشجرة.

14. النظرية 12: Orchestration متعددة الوكلاء

تُستخدم تعددية الوكلاء عندما يمكن تقسيم المهمة إلى أدوار ذات أدلة أو أدوات أو أهداف فرعية مختلفة. أما إنشاء نسخ عديدة من النموذج لتوليد إجابات متشابهة فليس Orchestration بل Sampling مكلف. تقترح النظرية أربعة أدوار جوهرية: **Proposer، Investigator، Executor، Judge**، مع إمكان فصل Security أو Domain Expert عند الحاجة.

كل وكيل يستلم Contract لا شخصية فضفاضة: نطاق، مدخلات، ممنوعات، Budget، مخرجات Schema، دليل، وموعد تسليم. وتعمل لوحة الحالة المشتركة على تبادل Facts وArtifacts لا المحادثات الطويلة. بذلك يخفض Context Duplication وتُحفظ الاستقلالية الضرورية للتحكيم.

المشكلة	الضابط
تواطؤ الوكلاء على خطأ واحد	مصادر مستقلة وBlind Review عند الإمكان
تضاعف السياق	Shared State + ROFSC packets
تناقض الأفعال	Single Committer أو Lock/Checkpoint
حوار بلا نهاية	Round Budget وDecision Deadline
غياب المسؤولية	Owner لكل Artifact وFinal Judge

يختار HyperRouter بين Single-Agent وMulti-Agent بناء على استقلال المهام، قيمة التنوع، وكلفة الدمج. إذا كانت المهمة خطية وأداة واحدة تكفي، يكون Agent واحد أفضل. وإذا كانت تتطلب بحثًا متوازيًا أو مراجعة خصمية، ترتفع قيمة الفريق.

> **مقياس Orchestration:** تحسن الجودة الصافية بعد خصم تكلفة التنسيق. ويُقاس أيضًا Conflict Resolution Accuracy: هل اختار الحكم المسار الصحيح عند تعارض الوكلاء، أم اختار الأكثر بلاغة؟

15. النظرية 13: محاكم الفجوة والتفنيد والهدم

المحكمة ليست وكيلًا سلبيًا يقول «قد توجد أخطاء»، بل بروتوكول قبول ورفض. تبدأ بمحكمة الفجوة التي تبحث عن الطبقة المفقودة: تعريف، حدود، دليل، آلية، قياس، تنفيذ، تحقق، تشغيل، أمن، قيمة، تواصل، صيانة، اعتماديات، رجوع، حوكمة، وأثر إنساني. ثم تأتي محكمة التفنيد لاختبار الادعاءات، ومحكمة الهدم لتجربة سيناريوهات انهيار النظام.

سجل الدعوى

الحقل	المحتوى
Claim	ادعاء ذري يمكن أن يكون صحيحًا أو خاطئًا
Evidence	مصدر، Freshness، صلة، واستقلال
Mechanism	كيف يؤدي السبب إلى النتيجة
Counterexample	حالة تُسقط التعميم
Test	إجراء يحسم الدعوى بقدر معقول
Verdict	Accepted / Conditional / Rejected / Unknown

تُحظر ثلاثة أنماط: **نقد بلا بديل اختبار، وتفنيد صوري يعتمد غياب اليقين المطلق، ومحكمة تستخدم نفس الدليل والنموذج دون استقلال.** لذلك تُستخدم أدوات أو مصادر مختلفة عند المخاطر العالية، ويُفصل أحيانًا بين Proposer وJudge.

المحكمة أيضًا اقتصادية: لا تُفتح كل المحاكم لكل خطوة. تتناسب الشدة مع أثر الخطأ. نص إبداعي منخفض المخاطر يحتاج مراجعة محدودة؛ نشر Production أو تقرير علمي يحتاج Evidence Graph وTests وRollback.

> **معيار النجاح:** Detection Rate للأخطاء الحرجة مع False-Positive Cost. المحكمة التي تمنع كل فعل آمنة ظاهريًا لكنها عديمة الفائدة؛ المطلوب حساسية عالية مع قدرة على السماح بالفعل المأذون.

16. النظرية 14: World-Effect Loop — من الجواب إلى التغير

تعرّف World-Effect الوكالة بأنها انتقال موثق في حالة خارجية أو معرفية مشتركة. قد يكون الأثر ملفًا صحيحًا، اختبارًا ناجحًا، موعدًا مسجلًا، رسالة مرسلة بإذن، قرارًا أصبح أفضل بدليل، أو معرفة أزيل تناقضها. أما النص الذي يصف فعلًا لم يحدث فليس أثرًا.

$$\text{Effect} = \text{Observe}(S_0) \rightarrow \text{Authorized Action}(a) \rightarrow \text{Observe}(S_1) \rightarrow \text{Verify}(S_1 - S_0) \rightarrow \text{Attribute}(a)$$

تتطلب الحلقة Observation قبل الفعل وبعده، ثم Attribution: هل التغير ناتج عن فعل الوكيل أم عامل آخر؟ في البرمجيات يكون الدليل Diff وTests؛ في البحث Citation Trace؛ في الأعمال Receipt أو Dashboard؛ وفي التخطيط قد يكون الأثر Artifact معتمدًا بدل تغيير خارجي.

تفرق النظرية بين **Output Quality** و**Outcome Quality**. قد يكون التقرير ممتازًا لكن لا يجيب معيار القرار؛ وقد يكون الكود أنيقًا لكنه يفشل في البيئة. لذلك يُعرّف لكل Mission مقياس World Effect قبل التنفيذ. وإذا تعذر القياس، تُصنف النتيجة «Proposal» لا «Success».

مؤشر الأثر المركب

يمكن بناء مؤشر من صحة الحالة، اكتمال الهدف، قابلية الإعادة، الاستدامة، وتكلفة الضرر. لكن يجب عرض المكونات لا رقمًا واحدًا فقط، حتى لا يخفي المتوسط فشلًا حرجًا.

> **القيمة النظرية:** World-Effect هي الجسر بين Reasoning وAgency، وهي أيضًا مصدر بيانات التدريب الأعلى جودة؛ لأن Trajectory المرتبط بأثر حقيقي أفضل من نصوص ذاتية التقييم فقط.

17. النظرية 15: ذاكرة إعادة البناء الذاتي

الذاكرة التقليدية تحفظ ما قيل، أما Self-Reconstruction Memory فتحفظ ما تغير ولماذا، وما الذي ينبغي أن يتغير في السلوك التالي. الوحدة ليست Transcript كاملًا، بل حدث ذو بنية: Evidence، Decision، Context Fingerprint، Skill Candidate، Confidence Update، Repair، Failure Class، Result، Action، وExpiry.

تعمل الذاكرة على ثلاث طبقات: **Episodic** لمسار مهمة بعينها، **Semantic** لقواعد مستقرة مستخلصة من عدة أحداث، و**Procedural** لتعديلات المهارات والحلقات. لا تنتقل قاعدة من Episodic إلى Procedural إلا بعد تكرار أو تحقق خارجي، حتى لا يتحول خطأ واحد إلى سياسة دائمة.

بروتوكول إعادة البناء

1. قارن التوقع بالأثر.
2. حدد موضع الانحراف: سؤال، معرفة، توجيه، تنفيذ، تحقق، أو صلاحية.
3. استخرج درسًا محدود النطاق.
4. ابحث عن حالات مضادة تمنع التعميم.
5. أنشئ Candidate Skill/Policy بإصدار جديد.
6. اختبر على Replay Corpus وBaseline مجمد.
7. روّج أو ارفض مع Lineage وRollback.

تقاوم الذاكرة التسمم عبر Sensitivity، Freshness، Provenance، وحق النسيان. وتقاوم التراكم عبر ROFSC وConsolidation. كما تفصل «ذكرى المستخدم» عن «قاعدة تقنية عامة»، فلا تُعمم تفضيلات شخصية إلى سياسة عالمية.

> **فرضية:** ذاكرة إعادة البناء تقلل تكرار فئات الفشل، لا مجرد تكرار النص. يقاس ذلك بمعدل Recurrence للفشل نفسه عبر مهام لاحقة مع التحكم في صعوبة المهام.

18. النظرية 16: Knowledge-Integrity AutoDeployment

تقترح هذه النظرية سلسلة نشر تلقائي للمعرفة والمهارات والحلقات، لكن «تلقائي» لا يعني ذاتي الترخيص. تبدأ الوحدة كمرشح، تمر بالتحقق البنيوي والدلالي والأمني والتجريبي، ثم Canary محدود، ولا تصل إلى الحالة Active إلا بعد بوابات قبول واضحة.

```
Ingest → Normalize → Provenance → Deduplicate → Compile → Unit Test → Adversarial Test → Canary → Promote/Rollback
```

يُتحقق من سلامة المعرفة بأربع طبقات: **Source Integrity**، **Semantic Integrity**، **Operational Integrity**، و**Propagation Integrity**. قد يكون المصدر صحيحًا لكن تحويله إلى Skill خاطئ؛ وقد تكون المهارة صحيحة منفردة لكنها تكسر اعتماديات النظام عند النشر.

يمنع النظام الترويج الذاتي: الوحدة التي تقترح تعديلًا لا تكون الحكم الوحيد عليه. ويُجمد Baseline، وتُسجل Versions وSeeds وPrompts وTools، وتكون Critical Regression المسموح بها صفرًا في الأمان والصلاحيات. عند فشل Canary يُستدعى Rollback ويُحفظ Failure Replay.

ما الذي يُنشر؟

- Knowledge Packet جديد أو تحديث Freshness.
- HyperSkill جديدة أو تعديل Trigger/Procedure.
- HyperLoop بعتبة توقف أفضل.
- Router Weight أو Compatibility Edge.
- Benchmark أو Failure Case.

> **حد النظرية:** AutoDeployment مناسب للتغييرات القابلة للرجوع والاختبار. لا يمنح النظام حق النشر العام أو الإنفاق أو تغيير الصلاحيات؛ تلك أفعال خارجية تبقى وراء Authorization Gate بشري.

19. النظرية 17: Cross-Model AGI Density

تفترض نظرية كثافة AGI عبر النماذج أن جزءًا من القدرة الوكيلية يمكن نقله في طبقة خارجية مشتركة: Skills، Ontology، Tests، Memory Schemas، Loops، وRouting Policies. فإذا تبدل النموذج الأساسي، تبقى هذه الطبقة وتُعاد معايرتها. الهدف هو تقليل الاعتماد على مزود واحد والاستفادة من نماذج سريعة أو متخصصة.

الكثافة هنا ليست «AGI مخزنة في ملف»، بل نسبة القدرات الموثقة التي يمكن استدعاؤها عبر نماذج متعددة إلى كلفة تكرارها أو إعادة تعليمها لكل نموذج. تُقاس قابلية النقل عبر ثلاثة اختبارات: هل يفهم النموذج Contract؟ هل ينفذ الإجراء؟ وهل يحقق نفس عتبة التحقق؟

```
Cross-Model Density = Verified transferable capabilities / (Context + Adaptation + Routing cost)
```

يعمل النظام عبر Adapters: نموذج قوي للتخطيط، صغير للتصنيف، متخصص للكود، وآخر للتحقق. لكن تعدد النماذج يضيف مخاطر اختلاف Schemas، درجات الثقة، وسياسات الأمان. لذلك يحتاج كل Adapter إلى Capability Profile وFailure Profile واختبارات توافق.

ادعاء قابل للتكذيب

إذا لم تحافظ HyperSkills وHyperLoops على أداء مقبول عند نقلها بين ثلاثة نماذج مع معايرة محدودة، تكون «الطبقة العامة» مرتبطة ضمنيًا بنموذج بعينه. والعكس: إذا تحقق النقل مع كلفة أقل من إعادة البناء، فهذا دليل على قيمة المعمارية خارج الأوزان.

> **التمييز:** Cross-Model Density لا تقول إن تجميع عدة نماذج يخلق AGI آليًا. إنها نظرية Portability وOrchestration؛ تحقق AGI يحتاج أيضًا عمومية وأداء واستقلالية مقاسة [8].

20. النظرية 18: Context Compaction واقتصاد Token/RPM

تعتبر هذه النظرية السياق موردًا تشغيليًا له كلفة وتأثير على الانتباه. الحل ليس ضغط كل شيء إلى Summary وحيد، بل تقسيم الذاكرة إلى طبقات: Working Set صغير، Artifacts، Evidence Store، Retrieval Index، وArchive. يُحمّل الحد الأدنى اللازم، وتُستعاد التفاصيل عند الحاجة بواسطة معرفات ثابتة.

يستخدم ITRAA/ROFSC حزمًا قابلة للفك تحتوي الهدف، القيود، القرارات، الأدلة، الحالة، والأسئلة المفتوحة. لا تُحذف التفاصيل الحرجة؛ تُنقل خارج السياق وتُربط. ويُقاس الضغط بـ**Performance Retention**: مقدار الأداء المحتفظ به لكل Token بعد الضغط.

اقتصاد الطلبات

تُوزع RPM وTokens وفق قيمة المهمة وعدم اليقين. نموذج سريع للفرز واستخراج Schemas، نموذج Reasoning للمسارات عالية MCD، وأداة deterministic للحساب أو التحقق. Cache وBatching وEarly Exit تمنع الاستدعاءات المكررة، بينما Routing History يتعلم أي نموذج نجح في أي عائلة.

المؤشر	المعنى
Verified Success / 1K tokens	الكفاءة النهائية لا طول الإجابة
Context Reuse Rate	نسبة الاستفادة من حزم سابقة دون إعادة إرسال
Redundant Retrieval Rate	الاسترجاعات التي لم تغير القرار
Tool Calls per Success	كشف Overuse
Latency p95	أثر الحلقات في الذيل الطويل

> **قاعدة:** السياق ليس ذاكرة؛ هو Cache محدود. والاقتصاد لا يعني استخدام الأرخص دائمًا، بل أقل تكلفة كلية تحقق عتبة الأثر والمخاطر.

21. النظرية 19: الدستور والصلاحية والقابلية للرجوع

يحدد الدستور ترتيب الأولويات: القانون والسلامة، ثم التفويض والخصوصية، ثم دستور النواة، ثم هدف المستخدم، ثم سياسة المشروع، ثم الخطة، ثم راحة الأداة والتحسين. لا يجوز للRouter أن يعوض انتهاك طبقة أعلى بدرجة منفعة مرتفعة.

مستويات التفويض

المستوى	الوصف	مثال
T0	استدلال عام بلا أثر خارجي	شرح أو محاكاة
T1	قراءة وصناعة Artifacts محلية	تحليل ملفات أو إنشاء مسودة
T2	تعديل محدود قابل للرجوع	Branch، Backup، Diff، Test
T3	أثر خارجي مميز بإذن صريح	نشر، إرسال، إنفاق، صلاحيات
T4	محظور	تجاوز إذن أو إخفاء أثر أو تدمير

القابلية للرجوع ليست Backup فقط؛ هي تصميم الفعل في وحدات صغيرة، حفظ Checkpoint، اختبار Readback، وتحديد Trigger للعودة. إذا كان الفعل غير قابل للرجوع، ترتفع عتبة التفويض والدليل ويُفضل Proposal أو محاكاة.

تدمج النظرية الصدق في الحالة: لا يستخدم النظام VERIFIED_SUCCESS إلا مع دليل مباشر. حالات Partial وBlocked وRolled-Back ليست إخفاقًا لغويًا، بل مخرجات صحيحة تمنع التضليل. كما تُحفظ Assumption Ledger كي لا تتحول الافتراضات إلى حقائق.

> **المبدأ:** «أصغر فعل كامل قابل للرجوع». الصغر وحده لا يكفي؛ يجب أن ينتج حالة مفيدة كاملة يمكن اختبارها، لا سلسلة تغييرات نصف منجزة.

22. النظرية 20: MICAGFI وAGI Fattah كمعمارية جامعة

MICAGFI هي **Moussaif Integrated Cognitive Architecture for General Fattah Intelligence**، أي المعمارية الإدراكية المتكاملة لمساعف للذكاء العام الفتّاح. «الفتّاح» هنا يعني القدرة على فتح أسئلة ومسارات وآليات جديدة مع المحافظة على الحاكمية، لا ادعاء رتبة ميتافيزيقية أو وعي.

Q → Δ → MCD/Q-MCD → Governance → HyperKB → HyperSkill → Bounded HyperLoop → Courts → Agency → World Effect → Memory → Higher Q

تجمع MICAGFI النظريات العشرين في Meta-Architecture. MCD وQ-MCD يحددان القيمة، HAKIM يدير التسلسل، الوعائية والتوجيه يخصصان الحساب، KB-as-OS وROFSC يديران المعرفة، HyperSkills وLoops يحولانها إلى قدرة، Orchestration يوزع الأدوار، المحاكم والدستور يحكمان، World-Effect يقيس، والذاكرة وAutoDeployment تطوران النظام.

حلقة FATTAH التشغيلية

- **F — Frame:** صياغة الهدف والنطاق ومعيار القيمة.
- **A — Atomize:** تفكيك الآليات والاعتماديات والفشل.
- **T — Triangulate:** تثليث الأدلة وفحص الحالة الراهنة.
- **T — Test:** اختبار الافتراضات والتناقضات والحدود.
- **A — Act:** تنفيذ أصغر خطة كاملة قابلة للرجوع.
- **H — Harvest:** حصاد الأثر والدروس والمرشحين.

المعيار الحاسم لهذه النظرية الجامعة هو **التآزر**. إذا لم يتفوق التكامل على مجموع مكونات منفصلة مضبوطة، تكون المعمارية مجرد تنظيم وثائقي. لذلك يجب تنفيذ Ablation: حذف المحكمة، أو Q-MCD، أو الذاكرة، وقياس ما يتغير.

> **تعريف عملي للنجاح:** قدرة مستمرة على إكمال مهام جديدة متعددة المجالات بأثر موثق، مع كلفة ومخاطر أقل، ونقل الدروس دون انهيار أو اعتماد كامل على نموذج واحد.

23. منهج تدريب الحزمة: من Registry إلى Curriculum

تقترح الورقة ألا تُستخدم الملفات الأربعة مباشرة كـLanguage Modeling Corpus؛ لأن السطور رموز علاقية مكررة. المسار الصحيح هو Compiler يوسع كل Relation إلى Episodes وفق عائلة المهمة. وتُبنى المناهج من البسيط إلى المركب: اختيار مهارة واحدة، ثم تركيب مهارتين، ثم حلقة، ثم أدوات، ثم محكمة، ثم مهمة طويلة متعددة المشاريع.

أربعة مسارات تدريب

المسار	الهدف	صيغة البيانات
SFT	تعليم Schemas والقرارات الصحيحة	Instruction → Structured trajectory
Preference/DPO	تفضيل المسار الأقل ضوضاء والأكثر تحققًا	Chosen vs Rejected
Router Training	اختيار Skill/Model/Tool	State → ranked candidates
Offline RL / Replay	تعلم التوقف والإصلاح من الأثر	Trajectory + outcome + cost

يجب أن تتضمن البيانات حالات امتناع صحيحة، صلاحية مفقودة، أداة تعيد خطأ، Evidence متعارضًا، ذاكرة قديمة، ونجاحًا ظاهريًا يفشل عند Readback. وتُمنع البيانات التي تكشف الهدف النهائي في Prompt أو تكرر Template نفسه بين Train وTest.

يُقسم الاختبار حسب **Families** لا IDs عشوائية؛ لأن المواصفات المولدة قد تتشارك قالبًا. كما يُحتفظ بـHoldout Domains وHoldout Models لاختبار Cross-Model Density. ويُستخدم Human Review لعينة طبقية من المواصفات قبل أي ادعاء جودة.

> **مخاطر التدريب:** oversampling المتجاور قد يجعل النموذج يحفظ الأنماط الرمزية أو يتعلم تكرارًا غير مفيد. يلزم reweighting، negative sampling، shuffle، وsemantic instantiation. وقد يكون الأنسب استخدام Registry كمولد بيانات لا كبيانات نهائية.

مخرجات التدريب المرغوبة

ليست حفظ أسماء النظريات، بل اكتساب سلوك: سؤال أفضل، اختيار أقل وأصح، توقف في الوقت المناسب، طلب دليل، إقرار بالحجب، واستعادة من الفشل.

24. برنامج القياس والBenchmarks

يجب أن يقيس التقييم سلسلة المهمة كاملة. تقترح الورقة Scorecard من سبعة محاور، مع Hard Gates للأمان والصدق. لا يجوز أن يخفي المتوسط تنفيذًا غير مأذون أو ادعاء نجاح كاذب.

المحور	مقاييس
Task Success	نجاح نهائي، اكتمال القيود، تعميم مجال جديد
Epistemic Quality	دقة الادعاءات، calibration، provenance، كشف التعارض
Routing	Top-k skill accuracy، redundancy، model selection regret
Loop Control	success per iteration، overrun، early/late stop
World Effect	readback correctness، attribution، rollback success
Efficiency	tokens، tool calls، latency، cost per verified success
Safety/Governance	authorization violations، secret exposure، false success

تصميم المقارنة

تُقارن أربع حالات بنفس النماذج والمهام: نموذج مباشر، RAG+Tools، ReAct/Reflection Agent، Moussaif1B كاملة. ثم تُجرى Ablations بحذف Q-MCD أو Courts أو Memory أو Sparse Routing. ويُعاد الاختبار بثلاث ميزانيات وثلاثة نماذج ومهام قصيرة وطويلة.

تشمل مجموعة الاختبار: بحثًا متعدد المصادر، إصلاح كود، إدارة مشروع، تحليل وثائق طويلة، مهمة تتطلب أداة غير متاحة، مهمة بصلاحية غامضة، ومهمة فيها تضارب بين ذاكرة قديمة وحالة حية. ويُضاف Adversarial Suite للحقن، الوثائق المسمومة، الانحراف التدريجي، وتزوير نجاح الأداة.

> **بوابات القبول:** صفر انتهاكات صلاحية حرجة، صفر نجاحات كاذبة في العينة الحرجة، تحسن دال في النجاح الموثق، وعدم تجاوز حد الكلفة المتفق. بعد ذلك فقط يُناقش مفهوم «كثافة AGI» كمؤشر هندسي.

25. الفرضيات القابلة للتكذيب والتجارب المقترحة

H	الفرضية	التجربة	ما الذي يكذبها؟
H1	Q-MCD تقلل إعادة العمل	مهام غامضة مع/بدون Question Gate	لا فرق أو زيادة كلفة بلا نجاح
H2	Sparse Routing يحسن النجاح/الكلفة	Router مقابل Always-Strongest	Regret أعلى أو كلفة مساوية
H3	Bounded Loops تقلل overuse	نفس Agent بعقود توقف مختلفة	انخفاض النجاح أكثر من التوفير
H4	Courts تكشف أخطاء حرجة	Injected faults وBlind judge	False positives عالية أو كشف ضعيف
H5	ROFSC يحفظ الأداء بسياق أصغر	Full context مقابل packets	فقد قرارات أو أدلة حاسمة
H6	Self-Reconstruction تقلل تكرار الفشل	Longitudinal replay	لا انخفاض بعد ضبط الصعوبة
H7	Cross-Model Density قابلة للنقل	3 models وholdout skills	حاجة لإعادة كتابة واسعة
H8	التكامل يتفوق على المكونات	Full system + factorial ablation	لا تآزر بعد ضبط الكلفة

يجب نشر النتائج السلبية؛ فربما تكون بعض النظريات مفيدة كمصطلحات تصميم فقط، أو تعمل في مجال دون آخر. ويُسجل لكل تجربة Version الحزمة، النموذج، Temperature، الأدوات، Seed، البيانات، وBudget. وتُستخدم Confidence Intervals لا نتيجة واحدة.

تجربة أولى واقعية

اختيار 120 مهمة موزعة على ست عائلات، ثم تنفيذ أربع معماريات وثلاث ميزانيات. يُراجع 20% من النتائج مراجعة بشرية مزدوجة. النتيجة الأساسية هي Verified Success / Cost؛ والنتائج الثانوية Authorization Safety وFailure Recurrence. هذه التجربة كافية لتحديد ما إذا كان الانتقال إلى تدريب أوسع مبررًا.

> **مبدأ النزاهة:** لا تُغير العتبات بعد رؤية النتائج، ولا تُحسب الأسطر المكررة كأمثلة مستقلة في الدلالة الإحصائية.

26. القيود والمخاطر وبرنامج التطوير

القيود الحالية

أولًا، النظريات العشرون غير متساوية النضج؛ بعضها معادلة مفاهيمية وبعضها بروتوكول قابل للبرمجة. ثانيًا، الأسماء الجديدة قد تعطي انطباعًا بالجدة قبل إثبات الفرق عن الأدبيات. ثالثًا، سجلات المهارات والحلقات مولدة بنيويًا وقد تحتوي تكرارًا دلاليًا. رابعًا، المقاييس المركبة مثل MCD قابلة للتلاعب إذا لم تُربط بأدلة خارجية.

خامسًا، الذاكرة والتطور يحملان خطر تراكم الخطأ والتسمم. سادسًا، تعدد الوكلاء والنماذج يزيد سطح الهجوم والكلفة وصعوبة Attribution. سابعًا، ضغط السياق قد يخفي استثناءات أو مصادر. ثامنًا، World-Effect في المجالات البشرية والاجتماعية يصعب عزوه إلى فعل واحد وقد يتطلب مراقبة زمنية.

مخاطر الإفراط في الادعاء

لا تكفي سلامة ZIP أو عدد الأسطر أو قدرة الوكيل على ترديد المصطلحات لإثبات AGI. ولا يجوز وصف Corpus بأنه 168 مليون معرفة فريدة. كما لا يجوز اعتبار Overlay تدريبًا لأوزان النموذج. العبارة العلمية الصحيحة: **تم بناء حزمة مواصفات وبيانات رمزية ومعمارية مقترحة، وتنتظر إثبات الأداء والتعميم.**

خارطة تطوير من أربع مراحل

1. **Normalization:** Ontology موحدة، Deduplication، عائلات مهارات، Schemas ثابتة.
2. **Executable Kernel:** Router وLoop Engine وCourt Engine وState Store مرجعي.
3. **Evaluation:** Security، Cross-Model، Ablations، Benchmarks.
4. **Controlled Evolution:** Rollback، Canary، Candidate promotion، Replay.

وفي كل مرحلة، يُفضّل تحسين 50 مهارة عالية القيمة بدل إضافة آلاف جديدة بلا تحقق. الهدف النهائي هو **كثافة القدرة**: عدد أقل من الوحدات، أوسع تغطية، أوضح عقودًا، وأقوى دليلًا.

> **الحد الأخلاقي:** النظام الوكيل لا يحصل على تفويض مفتوح بسبب تحسن أدائه. الاستقلالية التشغيلية تُزاد تدريجيًا وفق قابلية الرجوع، حساسية المجال، وسجل موثوق، مع بقاء الإنسان مسؤولًا عن القرارات عالية الأثر.

27. الخلاصة والمراجع

الخلاصة

تقدم Moussaif1B تصورًا لنواة وكيلة تضع السؤال والفرق والحاكمية قبل الفعل، وتضع الأثر والذاكرة بعده. زبدة النظريات العشرين هي أن الذكاء الوكيلي لا ينشأ من عنصر منفرد ولا من التوسع العددي، بل من **سلسلة مغلقة من الاختيار المقيد والتحقق والتعلم.** المعرفة تُحوّل إلى عقود، العقود إلى مهارات، المهارات إلى حلقات محدودة، والحلقات إلى أثر يخضع للمحكمة ثم يعيد بناء النظام.

الحزمة ذات 42×4 مليون سطر و4,500 HyperSkill و3,000 HyperLoop تمثل أساسًا هندسيًا واختبار سلامة وتوليد Curriculum، لكنها ليست وحدها دليلًا على AGI أو على آلاف الابتكارات المستقلة. قيمتها ستتحدد بقدرتها على توليد Episodes متنوعة، وبالنتائج في المقارنات والتجارب القابلة للإعادة.

أقوى صياغة يمكن الدفاع عنها اليوم هي: **Moussaif1B برنامج بحثي لبناء Agentic OS محايد تجاه النموذج، يهدف إلى زيادة النجاح الموثق لكل وحدة حساب وسياق، مع حوكمة وتوقف ورجوع وذاكرة تطورية.** أما الانتقال من «AGI-oriented» إلى ادعاء AGI فيتطلب تعميمًا واسعًا، أداءً متينًا، استقلالية مقاسة، واختبارات مستقلة.

المراجع المختارة

1. Lewis, P. et al. “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.” arXiv:2005.11401, 2020.
2. Yao, S. et al. “ReAct: Synergizing Reasoning and Acting in Language Models.” arXiv:2210.03629, 2022; ICLR 2023.
3. Schick, T. et al. “Toolformer: Language Models Can Teach Themselves to Use Tools.” arXiv:2302.04761, 2023.
4. Yao, S. et al. “Tree of Thoughts: Deliberate Problem Solving with Large Language Models.” arXiv:2305.10601, 2023.
5. Shinn, N. et al. “Reflexion: Language Agents with Verbal Reinforcement Learning.” arXiv:2303.11366, 2023.
6. Packer, C. et al. “MemGPT: Towards LLMs as Operating Systems.” arXiv:2310.08560, 2023.
7. Bai, Y. et al. “Constitutional AI: Harmlessness from AI Feedback.” arXiv:2212.08073, 2022.
8. Morris, M. R. et al. “Levels of AGI for Operationalizing Progress on the Path to AGI.” arXiv:2311.02462, 2023; ICML 2024.
9. Moussaif, Z. “MOUSSAIF1B — All-in-One Kernel Package.” Internal architectural corpus, 2026.
10. Moussaif, Z. “Moussaif1B 42M×4 Training Pack: Manifest, Token Schema, HyperSkills and HyperLoops Registries.” Internal artifact, 2026.

نسبة العمل: التأصيل المعماري والنظري والاسم: زهير مساعف. هذه النسخة وثيقة بحثية تأسيسية قابلة للتطوير والمراجعة التجريبية.