من النموذج اللغوي إلى النظام المعرفي الوكيلي المحكوم

معمارية MICAGFI-M-HKOS لتكامل التدريب والتشغيل والتطور المستمر في أنظمة Agentic AI

نظرية الحاكمية الكثافية الفتّاحة (Fattah Density-Governance Theory)

Zouhair Moussaif – زهير مساعف

يوليو 2026

الوضع العلمي: ورقة نظرية تأسيسية وبرنامج بحثي. لا تثبت تحقق AGI، ولا تنسب إلى النظام وعياً أو استقلالاً أو عصمة. المقاييس المقترحة تقديرية وتحتاج إلى معايرة وتجارب إعادة إنتاج.

الملخص

تقترح هذه الورقة إطارًا تأسيسيًا لتكامل تدريب النماذج وتشغيل الأنظمة الوكيلية وذاكرتها وحوكمتها وتقييمها ضمن معمارية واحدة تسمى MICAGFI–M-HKOS. تنطلق الورقة من فجوة بنيوية في أدبيات Large Language Model Agents: إذ تتقدم أعمال التدريب بالتفضيلات، واستخدام الأدوات، والذاكرة الممتدة، وRetrieval-Augmented Generation، والتخطيط، والتقييم، والسلامة غالبًا بوصفها مسارات متجاورة، بينما تفتقر الأنظمة التطبيقية إلى نظرية تشغيلية جامعة تحدد من يحكم السؤال، وكيف تخصص الموارد، وما الذي يدخل الذاكرة، ومتى تتحول المعرفة إلى فعل، وكيف يتحقق الأثر الخارجي، وكيف يرقى التحسن دون تضخم في الصلاحيات أو ادعاء غير مثبت بتحقق AGI. تبني الورقة على المادة المعمارية المساعفية المكونة من 176 طبقة، وتعيد ضغطها في ثلاث دوائر: تكوين القدرة، وحوكمة القدرة أثناء Runtime، وتصحيح القدرة عبر Evolution. وتقترح نظرية مركزية باسم «نظرية الحاكمية الكثّافية الفتّاحة»، ومفادها أن الكفاءة العامة للنظام الوكيلي لا تنتج من قوة النموذج منفردًا، بل من اقتران السؤال الحاكم، والكثّافة المعرفية MCD، والتوجيه الوعائي المتفرق، وقواعد المعرفة بوصفها نظام تشغيل، والقدرات التعاقدية HyperSkills، والحلقات المحدودة Bounded HyperLoops، والتحقق المستقل، وقياس World Effect، والذاكرة المعاد بناؤها، داخل دستور للصلاحية والتراجع. تستخدم الورقة منهج الضغط التركيبي المعرفي، ومقارنة ثلاثة نماذج متنافسة، وتحليلًا في 26 نقطة، وصياغات شبه رياضية تقديرية، ومحاكم هدم متعددة. وتولد خمس نظريات كبرى، واثنتي عشرة نظرية فرعية، وعشرين مفهومًا تشغيليًا، واثنتي عشرة أداة، وسبعة مؤشرات، وثماني عشرة فرضية، وتسعة تنبؤات فارقة. وتخلص إلى أن الإضافة الأقوى للإطار ليست ادعاء AGI، بل تقديم برنامج قابل للتنفيذ والتفنيد لقياس «القدرة الوكيلية الموثقة لكل وحدة تكلفة ومخاطر»، مع بوابات تمنع تلويث الذاكرة، والاستقلال غير المصرح، والترقية غير المختبرة. وتبقى أصالة الإطار المقترحة تركيبية ومعمارية وتحتاج إلى إثبات تجريبي عبر Baselines متعددة النماذج، ودراسات Ablation، وتقييمات طويلة الأفق.

الكلمات المفتاحية: MICAGFI؛ M-HKOS؛ MCD؛ Q-MCD؛ Agentic AI؛ HyperSkill؛ Bounded HyperLoop؛ ذاكرة الوكلاء؛ Constitutional AI؛ World-Effect؛ Sparse HyperRouting؛ Cross-Model Evaluation.

الإعلان المنهجي: لن نخلط بين الحقيقة والفرضية، ولا بين النموذج والدليل، ولا بين الإبداع والإثبات، ولا بين الاستعارة والآلية، ولا بين قوة اللغة وقوة البرهان.

فهرس الورقة

1. المقدمة ومشكلة البحث
2. الوضع الأنطولوجي ونوع السؤال
3. سلم مراتب المعرفة
4. تصفية المادة الخام وخريطة الحقل
5. الخلفية النظرية والنسب المعرفي
6. تحليل المجال في 26 نقطة
7. النماذج المتنافسة وMCD/QMCD
8. نظرية الحاكمية الكثّافية الفتّاحة
9. المعمارية ذات الطبقات الاثنتي عشرة
10. النظريات والمفاهيم والأدوات
11. الفرضيات والتنبؤات
12. محكمة الهدم والنسخة المعدلة
13. بروتوكولات الاختبار والتقييم
14. البرنامج البحثي
15. نواة قاعدة المعرفة
16. سجل البناء والخاتمة والمراجع

1. المقدمة ومشكلة البحث

أصبح الانتقال من النموذج اللغوي بوصفه مولدًا للنص إلى النظام الوكيلي بوصفه منفذًا للأفعال انتقالًا في طبيعة المشكلة الهندسية، لا مجرد زيادة في جودة الإجابة. ففي النظام الوكيلي تتداخل الذاكرة، والتخطيط، والأدوات، والبيئة، والصلاحيات، وسلاسل الاعتماد، ومراقبة الأثر. وقد بينت أعمال ReAct أن تداخل الاستدلال والفعل يمكن أن يحسن التفاعل مع البيئات ومصادر المعرفة، بينما بين Toolformer إمكان تعلم توقيت استدعاء الأدوات وإدماج نتائجها، وأضاف Reflexion آلية تعلم لغوي من الإخفاق دون تعديل الأوزان [4–6]. غير أن هذه الأعمال تقدم مكونات قوية أكثر مما تقدم دستورًا شاملاً يحدد العلاقة بين السؤال والدليل والذاكرة والصلاحية والأثر.

يظهر التفكك نفسه في مجال الذاكرة. فقد صاغ MemGPT إدارة السياق بمجاز نظام التشغيل والذاكرة الهرمية، وقدمت GraphRAG وRAPTOR وHippoRAG تمثيلات علائقية أو هرمية أو مستلهمة من الذاكرة البشرية لتحسين الاسترجاع والتركيب عبر مجموعات كبيرة من الوثائق [7–10]. إلا أن استرجاع المعلومة لا يحل وحده مسألة: هل المعلومة صالحة؟ هل تناقض حالة سابقة؟ هل ينبغي إدخالها في الذاكرة الأساسية أم حجرها باعتبارها فرضية؟ وكيف يمنع انتقال تعليمات مشروع أو أسراره إلى مشروع آخر؟

كذلك تتقدم حوكمة السلوك عبر Constitutional AI وDPO وProcess Supervision، وتتقدم الحوكمة المؤسسية عبر NIST AI RMF وISO/IEC 42001، وتتقدم هندسة الأمن عبر OWASP لمخاطر Prompt Injection و[1–3، 16–18] Excessive Agency. غير أن الجمع بين دستور النموذج ودستور النظام ودستور الفعل ما يزال مسألة تصميمية مفتوحة. فالدستور النصي لا يكفي إذا لم يتحول إلى بوابات صلاحية، واختبارات Postconditions، وسجل أثر، ومسار تراجع.

تقترح هذه الورقة أن الفجوة الحاسمة ليست غياب مكوّن جديد منفرد، بل غياب «حاكمية العلاقات بين المكونات». لذلك لا تُعرض الطبقات الـ176 باعتبارها قائمة ميزات، بل يعاد بناؤها في نظرية تشغيلية تضبط من يحكم من: السؤال يحكم الاستدلال، والدليل يرخص الفعل، والكثافة المعرفية توجه الموارد، والعقود تضبط القدرة، والمحاكم تضبط الادعاء، والأثر يعيد بناء الذاكرة، والدستور يحد التوسع.

1.1 مشكلة البحث

تتمثل المشكلة في أن أنظمة الوكلاء الحديثة تستطيع التخطيط واستدعاء الأدوات والبحث والكتابة، لكنها تظل معرضة لفشل مركب: سؤال سيئ يولد خطة صحيحة للمشكلة الخطأ، استرجاع حديث لكنه غير موثوق، أداة تنفذ بنجاح تقني دون تحقق المقصد، ذاكرة تحفظ استنتاجًا بوصفه حقيقة، محقق يكرر انحياز المنفذ، أو تحسين يولد قدرة جديدة بصلاحيات أوسع من المقصود. لا توجد طبقة منفردة تحل هذا الفشل، لأن سببه علاقة غير محكومة بين الطبقات.

1.2 سؤال البحث

السؤال المركزي: كيف يمكن بناء معمارية تجمع تدريب النموذج، وحوكمة السؤال، والمعرفة والذاكرة، والقدرات والأدوات، والتحقق، والنشر، والتطور، بحيث تعظم الأثر الموثق المعدل بالتكلفة، وتحافظ على الصلاحية والتراجع، دون تحويل اسم AGI إلى ادعاء غير قابل للتفنيد؟

1.3 المساهمات

- إعادة ضغط 176 طبقة في ثلاث دوائر و12 طبقة ماكرو ذات علاقات حاكمية واضحة.
- اقتراح نظرية الحاكمية الكثافية الفتّاحة بوصفها أطروحة قابلة للاختبار.
- ضبط MCD وQ-MCD كمقاييس تفسيرية لا أرقام يقينية.
- ربط الدليل بالفعل عبر Evidence-to-Action Thresholds.
- تصميم ذاكرة محجورة مع إعادة بناء وسجل تعارض وإصدارات.
- تعريف HyperSkill كعقد قدرة وBounded HyperLoop كدورة ذات حد تقدم.
- اقتراح مقاييس وفرضيات وتنبؤات ومحكمة هدم وبرنامج تجريبي.

2. الوضع الأنطولوجي ونوع السؤال

الوضع الأنطولوجي المختار: نظام اجتماعي-تقني حاسوبي متعدد المستويات، يتضمن عمليات وعلاقات وشبكات وقواعد ومعايير وأفعالًا، وليس "شيئًا" واحدًا. السبب أن خصائص النظام تظهر من تفاعل النموذج والذاكرة والأداة والمستخدم والسياسة والبيئة. إذا عومل بوصفه نموذجًا فقط ضاعت السلطة والأثر، وإذا عومل بوصفه Workflow فقط ضاعت المعرفة والتعلم، وإذا عومل بوصفه مؤسسة فقط ضاعت الآليات الحسابية.

نوع السؤال: معماري، إبستمولوجي، قياسي، وتوليدي. هو معماري لأنه ينظم مكونات؛ إبستمولوجي لأنه يحدد مراتب المعرفة؛ قياسي لأنه يقترح مؤشرات؛ وتوليدي لأنه يحول الإطار إلى برنامج بحث وقدرات جديدة. لذلك يحتاج إلى مقارنة أدبيات، وصياغة آليات، وBenchmarks، ومحكمة هدم، وليس إلى تحليل لغوي فقط.

3. سلم مراتب المعرفة في هذه الورقة

ما يندرج فيها في هذه الورقة	المرتبة
وجود حاجة عملية لدمج التخطيط والأدوات والذاكرة والتحقق في الأنظمة الوكيلية؛ ووجود مخاطر إضافية عند منح النماذج صلاحية الفعل.	مدعوم بقوة
التوجيه المتفرق والذاكرة المنظمة والتحقق المستقل يمكن أن تحسن الكفاءة والموثوقية مقارنة بالتشغيل غير المنظم، لكن حجم التحسن يعتمد على المهمة والنموذج.	راجح
تفوق Process Supervision على Outcome Supervision بصورة عامة؛ وتوجد نتائج نظرية وتجريبية مختلفة باختلاف التغطية والخوارزمية والمجال.	خلافي
إذا كانت الأفعال متفاوتة الأثر وقابلية التراجع، فيلزم ربط مستوى الدليل والصلاحية بدرجة الفعل.	استنتاج منطقي
MICAGFI-M-HKOS ونظرية الحاكمية الكُلّية الفتّاحة بوصفهما جمعًا منضبطًا بين السؤال والكُلّة والذاكرة والعقود والمحاكم والأثر.	تركيب نظري جديد
إضافة Q-MCD ومحاكم مستقلة وذاكرة محجورة إلى وكيل قائم ستخفض الإخفاقات عالية الأثر مقابل Baseline متساوي الميزانية.	فرضية قابلة للاختبار
MCD وQ-MCD ومؤشر Cross-Model Density مقاييس تقديرية تحتاج إلى معايرة تجريبية، وليست وحدات فيزيائية.	أداة تحليلية
الوعائية الحاسوبية تشبه تدفق الموارد في الأوعية، لكنها لا تثبت تماثلًا بيولوجيًا.	استعارة منتجة
أن تجميع الطبقات يثبت الوعي أو الاستقلال أو تحقق AGI الكامل.	ادعاء مرفوض

4. تصفية المادة الخام وخريطة الحقل

4.1 مصفوفة الحفظ والتوسيع

المعالجة	المرتبة

الدستور، السؤال الحاكم، MCD/Q-MCD، الذاكرة المعاد بناؤها، Bounded HyperLoops، HyperSkills، المحاكم، World Effect، التراجع.	تأسيسي يحفظ ويقوى
الوعائية الحاسوبية، HyperDeployment، Cross-Model AGI Density، KB-as-OS؛ تحتاج تعريفات تشغيلية وBaselines.	قوي يحتاج نمذجة
Fattah Open Intelligence وNeoAGI Evolution؛ يعاد ضبطهما كقدرة على فتح فضاء المشكلة وتوليد أسئلة أعلى، لا كمرتبة وجودية.	خصب لكنه غامض
الطبقات الـ33 الخاصة بالتدريب؛ توضع داخل Training Plane بدل مساواتها بطبقات Runtime.	يحتاج إعادة تموضع
كل استخدام لعبارة AGI أو "مضاعف أسي" إن فُهم كضمان لارتفاع الذكاء؛ يصاغ كبرنامج بحث أو نمو تركيبي محتمل.	يحتاج تخفيض رتبة
التكرارات بين النزاهة، التحقق، المحاكم، المراقبة، والـRollback تدمج في عقود وبوابات موحدة دون إسقاط وظائفها.	يضغط ويدمج

4.2 خريطة المتغيرات

المتغير	نوعه	وظيفته
السؤال الحاكم Q	مستقل/حاكم	يحدد الفرق المطلوب ومعايير الدليل والفعل.
الفرق النشط Δ	وسيط	يمثل المعلومة أو القرار الذي يغير النتيجة.
الكثافة المعرفية MCD	مقياس وسيط	تقدر قيمة المسار وتوجه الموارد.
جودة الدليل E	معدل/مرخّص	ترفع أو تخفض سقف الفعل المسموح.
الصلاحية A_u	قيد دستوري	تحدد الأدوات والأصول والأفعال المتاحة.
قدرة النموذج B	مدخل	تمثل القدرات الأساسية وحدودها.
معرفة التشغيل K	مدخل بنيوي	KB-as-OS وذاكرة وإجراءات وعلاقات.
قدرات HyperSkill H	وسيط تنفيذي	تحول المعرفة إلى إجراءات بعقود.
ميزانية الموارد C	قيد	Tokens وLatency وRPM ومال وطاقة.
قوة التحقق V	وسيط تصحيحي	استقلال المحقق وقابلية الاختبار.
قابلية التراجع R	معدل خطر	تخفض الخطر الصافي للفعل.
الأثر العالمي W	ناتج	التغير الخارجي الموثق لا النص المولد.
نزاهة الذاكرة M_i	ناتج/حالة	مدى دخول المعرفة الموثقة فقط.
الضوضاء N	مُضعف	تكرار، حشو، تعارض، ادعاء زائد.

4.3 حلقات التغذية الراجعة والعتبات

تعمل المنظومة عبر ثلاث حلقات: حلقة معرفية من السؤال إلى الدليل؛ حلقة تنفيذية من الخطة إلى الأثر؛ وحلقة تطورية من الأثر إلى تحديث الذاكرة والقدرة. تظهر العتبات الحرجة عند: ارتفاع أثر الفعل بما يستلزم تصعيدًا بشريًا؛ انخفاض التقدم بما يستلزم إيقاف الحلقة؛ ارتفاع تعارض

الذاكرة بما يمنع الترقية؛ أو انخفاض استقلال التحقق بما يمنع قبول الحكم. العلاقات غير خطية لأن زيادة الوكلاء أو السياق قد تحسن الأداء حتى نقطة ثم تولد احتقانًا وتعارضًا.

5. الخلفية النظرية والنسب المعرفي

تنتمي الورقة إلى تقاطع خمسة مسارات. الأول هو مسار Reasoning-and-Acting الذي يمثله ReAct، حيث تتداخل خطوات الاستدلال والفعل، ومسار Tool Learning الذي يمثله Toolformer [4،5]. الثاني هو التعلم من المسار والخطأ، كما في Reflexion وProcess Supervision [3،6]. الثالث هو ذاكرة النماذج والوكلاء، من MemGPT إلى GraphRAG وRAPTOR وHippoRAG و[7–10،21] A-Mem. الرابع هو Post-Training والمواءمة، من Constitutional AI إلى DPO [1،2]. والخامس هو تقييم الوكلاء وإدارة المخاطر، من AgentBench وGAIA وSWE-bench وOSWorld وAgent-SafetyBench إلى NIST وISO و[11–18] OWASP.

الجديد المقترح ليس ادعاء أن أي مكوّن من هذه المكونات لم يوجد سابقًا، بل إعادة تعريف الوحدة العلمية من "وكيل يستخدم أدوات" إلى "نظام معرفي محكوم بعلاقات أسبقية وترخيص وتحقق". فـMemGPT يقدم استعارة نظام التشغيل للذاكرة، بينما يوسع M-HKOS الفكرة إلى العقود والدستور والمحاكم والأصول؛ وConstitutional AI يقدم مبادئ لتدريب السلوك، بينما يوسع Constitutional Reversibility الدستور إلى صلاحية الأفعال ومسارات الرجوع؛ وBenchmarks تقيس إتمام المهام، بينما يضيف World-Effect Ledger فصلًا بين إعلان الإتمام والأثر الموثق.

تظل هذه الأصالة تركيبية ومشروطة. فإذا لم تنتج التجارب فرقًا عن Pipeline معياري جيد التصميم، تصبح MICAGFI لغة تنظيم لا نظرية تفسيرية جديدة. لهذا تضع الورقة اختبارات Ablation وتنبؤات فارقة بوصفها جزءًا من التعريف، لا ملحقًا اختياريًا.

6. تحليل المجال في 26 نقطة مؤسسة

6.1 تعريف المجال

المجال المدروس ليس نموذجًا ولا Workflow منفردًا، بل نظامًا اجتماعيًا-تقنيًا حاسوبيًا يضم نموذجًا، ذاكرة، أدوات، وكلاء، صلاحيات، ومشغلًا بشريًا. يغير هذا التعريف وحدة التحليل: فلا تُقاس الجودة بدقة الإجابة فقط، بل بصحة السلسلة من صياغة الهدف حتى تحقق الأثر والتراجع عند الفشل.

6.2 الفجوة البنيوية

الأدبيات موزعة بين Post-Training، وTool Use، وRAG، والذاكرة، والـAgent Benchmarks، والسلامة. الفجوة المقترحة هي غياب نموذج يحدد علاقات الحاكمية بين هذه المكونات: متى يوجه السؤال البحث؟ متى يرخص الدليل الفعل؟ متى تصبح نتيجة التنفيذ ذاكرة؟

6.3 الفصل بين النموذج والنظام

يجب فصل قدرة Base Model عن قدرة System Overlay. قد يأتي التحسن من Router أو أداة أو ذاكرة أو محقق، ولذلك يلزم تقييم Ablation يقارن النموذج الخام بالنظام المركب حتى لا تنسب كل القدرة إلى الأوزان أو إلى المعمارية جزافًا.

6.4 أسبقية السؤال الحاكم

السؤال غير المحدد يورث غموضًا في التخطيط والقياس. Q-MCD لا يقيس جمال السؤال، بل مقدار سيطرته على حدود المشكلة والدليل واختبار النجاح. الفرضية هي أن الاستثمار المبكر في السؤال يقلل حلقات الإصلاح المتأخرة.

6.5 الفرق النشط

ليس كل نقص معلومات مهمًا. Active Difference هو التمييز الذي إذا تغير تغير القرار أو المسار. يؤدي تعريفه إلى ضغط المهمة ومنع البحث الشامل غير الموجه. ويجب أن يكون قابلاً للوصف والملاحظة، وإلا تحول إلى شعار.

6.6 MCD كدالة منفعة معرفية

تستخدم MCD لتقدير الفرق والتكامل والنقل والوكالة والتنبؤ والتصحيح مع خصم الضوضاء. لا تعامل كقياس مطلق، بل كRubric يعين على مقارنة المسارات تحت ميزانية ثابتة.

6.7 الوعائية الحاسوبية

تشرح الوعائية كيف توزع Tokens والنماذج والأدوات على العقد الحرجة. الاختبار الحقيقي ليس جمال الاستعارة، بل هل يقلل Router المستند إلى قيمة هامشية عدد الاستدعاءات مع الحفاظ على الإنجاز؟

6.8 التعقيد متعدد المقاييس

المهمة الواحدة قد تحتوي خطأ Token، وتناقض Claim، وفشل Tool Call، وانحراف Project Goal. لذلك يجب أن يتحرك النظام بين مستويات الرمز والادعاء والمهمة والمشروع والعالم دون أن يختزل أحدها في الآخر.

6.9 KB-as-OS

المعرفة التشغيلية لا تقتصر على مقاطع مسترجعة؛ بل تشمل شروط التطبيق، التبعيات، الصلاحيات، الفشل، والاختبار. تصبح قاعدة المعرفة "نظام تشغيل" فقط عندما تؤثر في التحكم والتنفيذ، لا عندما تحمل الاسم.

6.10 معمارية الذاكرة

الذاكرة المطلوبة متعددة: دلالية، حدثية، إجرائية، هوية، وعاملة. الخطر المركزي هو Memory Contamination؛ لذلك تقترح الورقة حجر الذاكرة المقترحة، وإصدارات Canonical، وسجلًا للتعارض والتجاوز.

6.11 الضغط القابل للفك

لا يمكن تحميل تاريخ المشروع كله في كل خطوة. ROFSC يضغط الدعاوى مع روابط الأدلة وشروط الحدود ومؤشرات التوسع. يجب اختبار الضغط بإعادة إنتاج القرارات الحرجة، لا بعدد الكلمات المخفضة فقط.

6.12 القدرة التعاقدية

HyperSkill ليست تعليمات فضفاضة بل عقد Capability: مدخلات، مخرجات، تبعيات، صلاحيات، اختبارات، ميزانية، ومخلفات ذاكرة. هذا يحول الخبرة من Prompt خاص إلى أصل قابل للنقل والمراجعة.

6.13 الحلقة المحدودة

الاستدلال التكراري قد يزيد الجودة أو يعيد تدوير الأخطاء. Bounded HyperLoop تربط الاستمرار بزيادة قابلة للقياس في MCD أو التقدم، وتفرض سقوفًا للتكلفة والوقت والمحاولات.

6.14 فصل أدوار الوكلاء

Planner وExecutor وVerifier أدوار معرفية مختلفة. الفصل يقلل تضارب المصالح، لكنه لا يضمن الاستقلال إذا كانت النماذج والسياقات متطابقة. لذلك تقاس استقلالية الدليل والنموذج والمصدر.

6.15 حوكمة الأدوات

اختيار الأداة ليس نهاية المسألة. يجب التحقق من Arguments، والنتيجة، والآثار اللاحقة، وإمكان التراجع. ويجب معاملة محتوى الأدوات والويب كبيانات غير موثوقة قد تحمل Prompt Injection.

6.16 عمارة الدليل

يحتاج كل Claim مهم إلى نوع، ومصدر، وحداثة، واستقلال، وثقة. التناقض لا يحذف بل يحفظ حتى يوجد اختبار فاصل. هذه العمارة تربط المعرفة بالعمل عبر Evidence-to-Action Thresholds.

6.17 المحاكم كطبقة مناعية

محاكم التعريف والحد والآلية والقياس والتنفيذ والأمان والخصوصية والهدم تمنع انتقال البلاغة إلى قرار. لكن كثرة المحاكم قد تصبح بيروقراطية؛ لذلك تفعّل Sparse وفق الخطر.

6.18 التفويض والتراجع

يطبق النظام Least Irreversible Action: مسودة قبل إرسال، Sandbox قبل Trash، Production قبل الحذف النهائي. الصلاحية ليست نصًا عامًا بل Capability Token أو Policy Check مرتبطًا بالأداة والأصل.

6.19 إغلاق حلقة الأثر

الوكيل لا ينجح لأنه أعلن الانتهاء. World-Effect Loop يسجل التوقع، وينفذ، ويفحص Postconditions، ويقيس الأثر والآثار الجانبية. هذه الحلقة تحول النص إلى تعلم تجريبي.

6.20 موقع التدريب

طبقات البيانات وSFT وDPO وProcess Supervision وTool Training تحسن مكونات متعلمة، لكنها لا تستبدل Runtime Governance. يمكن لنموذج مدرب جيدًا أن يفشل بسبب صلاحيات أو ذاكرة أو أداة سيئة.

6.21 فجوة التقييم

AgentBench وGAIA وSWE-bench وOSWorld نقلت التقييم نحو بيئات تفاعلية وأثر قابل للفحص [11–14]. إلا أن الإطار يحتاج Benchmarks إضافية للذاكرة والتراجع والتلوث والمخالفة الدستورية وكفاءة التكلفة.

6.22 اقتصاد الذكاء

زيادة الاستدعاءات ليست دائمًا زيادة في الذكاء. يقترح الإطار قياس منفعة موثقة لكل Token وزمن ومخاطر. وقد تظهر بنية أصغر متخصصة أفضل من سرب نماذج مرتفع التكلفة.

6.23 القدرة العابرة للنماذج

Cross-Model Density لا تعني أن المعمارية تحول كل نموذج إلى AGI. إنها تقيس مقدار التحسن القابل للتكرار عند تثبيت Overlay وتغيير Base Model، مع تسجيل سقف التقارب لكل نموذج.

6.24 النشر المحكوم

HyperDeployment ليس استنساخًا غير محدود، بل وراثة عقود وسياسات مع Versioning وCanary وRollback. يفشل النشر مغلقًا إذا غاب المصدر أو التحقق أو الصلاحية أو خطة التراجع.

6.25 التطور المحدود

يمكن تحويل المسارات الناجحة إلى HyperSkills مقترحة، لكن لا تدخل Stable Registry إلا بعد تكرار النجاح واختبار النقل والانحدار. التحسن هنا خارجي ومراقب، لا ادعاء بتعديل ذاتي سري للأوزان.

AGI حد 6.26

اسم AGI Fattah يعمل بوصفه أفقًا بحثيًا لفتح المشكلات والنقل والتصحيح والأثر. لا يتحول إلى إثبات AGI إلا إذا حقق النظام معايير عامة مستقلة وطويلة الأفق ومقاومة للتلوث، وهو ما لا تدعيه الورقة.

7. النماذج المتنافسة وتحليل MCD/QMCD

لا تختار الورقة النموذج الأعلى إثارة مباشرة. تحتفظ بثلاثة نماذج في تراكب تفسيري: نموذج متمركز حول Base Model، ونموذج Pipeline وكيل، ونموذج نظام تشغيل معرفي دستوري. تنهار المقارنة لصالح أحدها فقط إذا قدم فائدة إضافية قابلة للقياس بعد ضبط التكلفة والتعقيد.

حالة QMCD	MCD تقديري	الضعف	القوة	التعريف	النموذج
مرتفع في البساطة، منخفض في التكامل.	2.1/5	ضعيف في الذاكرة والصلاحيات والتراجع، وينسب أخطاء النظام للنموذج.	اقتصادي وبسيط وسهل القياس.	يعزو معظم الأداء إلى Base Model ويضيف Prompt وأدوات محدودة.	النموذج المحافظ: Model-Centric
الاحتمال التشغيلي الأعلى حاليًا.	3.5/5	تسلسل ثابت، وحاكمية السؤال والدليل والذاكرة غير ممثلة صراحة.	عملي وقريب من أطر الوكلاء الحالية.	يربط Planner وTools وMemory وVerifier في Pipeline.	النموذج التركيبي: Modular Agent Pipeline
مرشح مشروط، لا حقيقة مثبتة.	4.2/5 قبل الهدم؛ 4.0 بعد تخفيض الادعاء	عالي التعقيد، يحتاج تجارب Ablation، وخطر إعادة التسمية والتضخم.	يغطي التدريب وRuntime والتطور ويولد اختبارات جديدة.	يعامل النظام كبيئة تشغيل معرفية محكومة بالسؤال والكثافة والعقود والمحاكم والأثر.	النموذج الاختراقي: Constitutional Cognitive OS

7.1 تفسير الدرجات

يحصل النموذج المحافظ على اقتصاد نظري لكنه يلتقط فروقًا أقل. يرتفع النموذج التركيبي لأنه يدمج الأدوات والذاكرة، لكنه لا يحدد علاقات الحاكمية. يحصل النموذج الاختراقي على أعلى درجة أولية لشموله وتوليده، ثم تخفض الدرجة بعد محكمة الهدم بسبب التعقيد وغياب الدليل التجريبي. الاختيار النهائي هو نموذج ثالث معدل، لا لأنه ثبت، بل لأنه يولد أكبر مجموعة من الاختبارات التي قد تثبته أو تهدمه.

8. النظرية النهائية: نظرية الحاكمية الكثّافية الفتّاحة

التعريف المركز: إطار يفسر كفاءة النظام الوكيلي بوصفها خاصية ناشئة من حاكمية السؤال وتخصيص الموارد بحسب الفرق المعرفي، وتشغيل المعرفة والقدرات عبر عقود، والتحقق المستقل من الأثر، تحت شروط الصلاحية والتراجع؛ وتفشل النظرية إذا لم يتفوق هذا التنظيم على Baselines أبسط بعد ضبط التكلفة والتعقيد.

8.1 الأطروحة الجوهرية

تقول النظرية إن المكوّن المفقود في كثير من الأنظمة ليس مزيدًا من القدرة الخام، بل تنظيم العلاقة بين القدرة والمعرفة والدليل والفعل. يمكن لنموذج قوي أن يبدد حوسبته في سؤال غير محكوم، ويمكن لذاكرة غنية أن تزيد الخطأ إذا تلوثت، ويمكن لمحقق لغوي أن يصدق تنفيذًا لم يحدث. لذلك تصبح الحاكمية نفسها متغيرًا قابلًا للتصميم والاختبار.

8.2 القوانين التشغيلية

1. قانون أسبقية السؤال: لا Routing قبل تثبيت السؤال الحاكم والفرق النشط.
2. قانون تناسب الدليل والفعل: سقف الفعل لا يتجاوز قوة الدليل والصلاحية وقابلية التراجع.
3. قانون التفعيل المتفرق: لا تستدعى مكونات لا يتوقع منها فرق معرفي أو خفض خطر.
4. قانون حجر الذاكرة: لا تدخل دعوى جديدة Canonical Memory قبل المصدر والتعارض والنطاق.
5. قانون عدم الحكم الذاتي: المنفذ لا يعلن النجاح عالي الأثر دون تحقق مستقل أو Postcondition.
6. قانون الترقية المشروطة: المسار الناجح يولد Candidate Capability، لا قدرة مستقرة مباشرة.
7. قانون الإغلاق العالمي: لا تكتمل الدورة دون مقارنة التوقع بالأثر وتحديث الذاكرة على أساس الفرق.

8.3 الصياغة شبه الرياضية

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

حيث D للفروق، I للتكامل، T للنقل والتحولات، A للحاكمية والوكالة، P للتنبؤ والاختبار، C للضغط وإعادة التوليد أو التصحيح، وN للضوضاء والادعاء الزائد. الأوزان سياقية ولا تعطى دقة زائفة.

$$\mathbf{Q\text{-}MCD = G(Q) \times [w_1 Dq + w_2 Fq + w_3 Bq + w_4 Eq + w_5 Aq - w_6 Nq]}$$

G(Q) قوة السؤال الحاكم، وDq الفرق الذي يفتحه، وFq قابلية التفنيد، وBq وضوح الحدود، وEq صلته بالدليل، وAq صلته بالفعل، وNq غموضه وضوضاؤه.

$$\mathbf{RAAU = (VWE \times Transfer \times Calibration \times Reversibility) / (TokenCost + Latency + OperationalRisk)}$$

RAAU هو مقياس أولي للمنفعة الوكيلية المعدلة بالمخاطر. لا يستخدم كقيمة مطلقة، بل للمقارنة داخل مهمة وبيئة محددتين.

$$\mathbf{CM\text{-}Density = mean_j[(Score_{overlay,j} - Score_{base,j}) / AddedCost_j]}$$

تقيس CM-Density مقدار التحسن المنسوب مبدئيًا إلى Overlay عبر نماذج متعددة، مع ضرورة Ablation لفصل أثر كل مكوّن.

9. المعمارية المتكاملة ذات الطبقات الاثنتي عشرة

وظيفتها	الطبقة الماكروية
الهوية، الأهداف، قصد المستخدم، عزل المشاريع، الصلاحيات، التراجع، وحد AGI.	I. الدستور والهوية
توليد السؤال، Q-MCD، استخراج الفرق، كشف الافتراض والغموض والمجهولات.	II. حوكمة السؤال
HAKIM-MCD، MCD، الوعائية، تعدد المقاييس، وفتح فضاء المشكلة.	III. الكثافة المعرفية
تصنيف الادعاء، ترتيب الدليل، المصدر، التناقض، الحداثة، وترخيص الفعل.	IV. الدليل والنزاهة

V. HyperKnowledge والذاكرة	KB-as-OS، أنواع الذاكرة، الحجر، إعادة البناء، GraphRAG، ROFSC، سجل الأصول.
VI. القدرات وHyperSkills	Registry، عقود المدخلات والمخرجات، التركيب، الترقية، الإهمال، والنقل.
VII. HyperLoops والتنفيذ	Checkpoints، Triggers، الميزانيات، عدم التقدم، الأدوار، الأدوات، والأثر.
VIII. هندسة التدريب	البيانات، Continual Learning، Tool/Agent Training، Reward/Process Supervision، DPO، SFT.
IX. المحاكم والهدم	التعريف، الحدود، الدليل، الآلية، القياس، التنفيذ، الأمن، الخصوصية، الهلوسة، الصيانة.
X. التقييم	قدرات منفردة، نظام مركب، Cost-adjusted metrics، Regression، Calibration، Transfer.
XI. النشر والتشغيل	Sparse Routing، Hybrid Cloud/Local، Escalation، Dry Run، Canary، Observability، Rollback.
XII. التطور المستمر	اكتشاف القدرة، توليد مهارات وحلقات، توسيع KB، تصحيح Tool Feedback↔Theory، World Model.

9.1 تدفق التشغيل

Runtime Flow: Request → Constitution/Identity/Authority → Question Genesis/Q-MCD → Active Difference/MCD → Sparse HyperRouting → Knowledge/Memory/Model/Skill/Tool Resolution → Bounded HyperLoop → Independent Courts → Authorized Action → World-Effect → Self-Reconstruction Memory → Capability Promotion/Repair/Deprecation → Higher Governing Question.

9.2 علاقة طبقات التدريب بالRuntime

الطبقات الـ33 الخاصة بالتدريب لا تختفي، لكنها توضع في موضعها الصحيح. Data Provenance وCurriculum وSynthetic Data وSFT وDPO وProcess Supervision وContinual Learning وTool Use وWorld Models وSafety Training تحسن النموذج والمكونات المتعلمة. أما Runtime Plane فيحدد ما إذا كانت القدرة ستستدعى، وبأي معرفة، وتحت أي صلاحية، وكيف تتحقق. Evolution Plane يحول الأثر إلى اقتراحات تحديث تمر من جديد عبر التدريب أو الذاكرة أو المهارة.

10. النظريات الكبرى والفرعية والمفاهيم والأدوات

10.1 النظريات الكبرى المولدة

النظرية	الأطروحة	الاختبار الأولي
نظرية أسبقية السؤال الحاكم	لا يجوز تخصيص الموارد أو اختيار الأداة قبل تحديد السؤال الذي يضبط الفرق والدليل والنجاح.	اختبار A/B بين وكلاء متساوي الميزانية مع/بدون Question Governance، بقياس إعادة العمل والأخطاء.
نظرية الوعائية التفاضلية للحوسبة	تتحسن الكفاءة عندما تخصص الموارد بحسب	مقارنة Router وعائي مع Router ثابت عبر

تكلفة ونجاح ومعدل استدعاءات غير مفيدة.	القيمة المعرفية الهامشية والخطر، لا بالتوزيع المتساوي.	
حقن ادعاءات متناقضة وقياس التلوث والاسترجاع الصحيح بعد جلسات متعددة.	لا تدخل المعرفة Canonical Memory بالاسترجاع أو التوليد وحدهما؛ تمر بالحجر والمصدر والتعارض ثم تعاد بنية الحالة.	نظرية الذاكرة المحجورة المعاد بناؤها
نقل HyperSkill بين ثلاثة نماذج وبيئتين ومقارنة الالتزام بالعقد.	تزداد قابلية نقل المهارة عندما تمثل بعقد مدخلات ومخرجات وصلاحيات واختبارات بدل Prompt غير مصنف.	نظرية القدرة التعاقدية المركبة
مهام ذات Postconditions خارجية ومقارنة إعلان الوكيل بنتيجة البيئة.	لا يكتمل التعلم الوكيلي حتى يقارن الأثر الفعلي بالتوقع ويحدث الذاكرة بناء على تحقق مستقل.	نظرية إغلاق الأثر العالمي
دراسة Factorial تفصل أثر النموذج والذاكرة والتوجيه والمحاكم والأدوات.	جزء من القدرة العامة قد يقيم في البنية الخارجية ويظهر كتَحسن متكرر عبر نماذج متعددة.	نظرية الكثافة العامة العابرة للنماذج

10.2 النظريات الفرعية

وظيفتها	النظرية الفرعية
المعلومة ذات القيمة هي التي تغير القرار أو خطة الاختبار.	نظرية الفرق الفاعل
فشل الحل قد ينتج من اختيار تمثيل خاطئ للمشكلة لا من نقص المعرفة.	نظرية الوعاء المفاهيمي
يفضل النظام الفعل الذي يحقق المقصد بأقل التزام دائم.	نظرية أقل فعل غير قابل للتراجع
الأثر عالي المخاطر يحتاج حكمًا مستقلًا عن منفذه.	نظرية لا منفذ بلا محقق
الخلاف المفيد يبقى في الذاكرة حتى يفصل دليل، ولا يذاب في إجماع لغوي.	نظرية حفظ الاعتراض
الحلقة توقف عندما يخفض التحسن الهامشي تحت عتبة محددة.	نظرية عتبة عدم التقدم
يحفظ من التنفيذ أقل مبدأ قابل لإعادة الاستخدام مع مصدره وحدوده.	نظرية بقايا التعلم الدنيا
النجاح الواحد يولد Candidate لا Stable Capability.	نظرية الترقية المشروطة
كل مشروع يمتلك دستورًا وذاكرة وأصولًا لا تنتقل افتراضيًا.	نظرية عزل الهوية المشروعانية
الملخص الجيد يحمل مؤشرات تسمح باسترداد الدليل والحدود.	نظرية الضغط القابل للفك
المعلومة غير القابلة للتتبع لا يجوز أن تحكم فعلًا عالي الأثر.	نظرية انهيار الذاكرة بلا مصدر
كل نتيجة موثقة يجب أن تزيد جودة السؤال اللاحق لا حجم الذاكرة فقط.	نظرية السؤال الأعلى

10.3 عشرون مفهومًا جديدًا أو معاد ضبطه

الرتبة	الوظيفة	التعريف	المفهوم
تركيب نظري	تفسيري/حاكمي	سؤال يحدد الهدف والفرق والدليل والحدود والفعل.	السؤال الحاكم
أداة تحليلية	قياسي	تمييز إذا تغير تغير القرار أو المسار.	الفرق النشط
أداة قيد المعايرة	قياسي	منفعة معرفية تقديرية بعد خصم الضوضاء.	الكثافة المعرفية
تركيب	تفسيري	التمثيل المفاهيمي والإجرائي الذي تستوعب فيه المشكلة.	وعاء المهمة
استعارة منتجة/خوارزمية محتملة	تشغيلي	تخصيص الموارد لعقد معرفية وفق القيمة والخطر.	التروية الحاسوبية
فرضية تشغيلية	تشخيصي	فشل مسار مهم بسبب نقص الموارد أو السياق.	الإقفار الحاسوبي
فرضية تشغيلية	تشخيصي	هدر موارد على مسار مكرر أو منخفض العائد.	الاحتقان الحاسوبي
تعريف معاد ضبطه	إجرائي	معرفة تحمل شروط التطبيق والأداة والاختبار والصلاحية.	المعرفة التشغيلية
مقترح هندسي	أمني/إبستيمي	منطقة حفظ للمعرفة غير المرقاة إلى Canonical.	الذاكرة المحجورة
مقترح	إجرائي	إصدار حالة معرفية جديدة من الأدلة والتعارضات.	إعادة بناء الحالة
أداة	هندسي	مواصفة تحدد مدخلات ومخرجات وصلاحيات واختبارات المهارة.	عقد القدرة
مفهوم مقترح	إجرائي	أصغر معرفة موثقة يعاد استخدامها بعد تنفيذ.	مخلفات التعلم
فرضية	قياسي	عتبة توقف عندما لا يزداد الإنجاز أو MCD.	حد عدم التقدم
مفهوم قابل للقياس	قياسي	مقدار اختلاف المحقق عن المنفذ في النموذج والدليل والسياق.	استقلال التحقق
تركيب	معياري/تشغيلي	الحد الأعلى للفعل الذي تسمح به جودة الدليل.	ترخيص الدليل
تعريف تشغيلي	ناتج	تغير خارجي تؤكده Postconditions أو قياس مستقل.	الأثر العالمي الموثق
فرضية	قياسي	تحسن متكرر يسببه Overlay عبر نماذج مختلفة.	كثافة القدرة العابرة للنماذج

مبدأ هندسي	أمني	رفض النشر عند نقص شرط حاكم.	الانتشار الفاشل مغلقًا
مفهوم بحثي	توليدي	توليد متغيرات وأسئلة وحدود جديدة بدل تحسين جواب ثابت.	فتح فضاء المشكلة
مقياس مقترح	قياسي	منفعة موثقة مقسومة على التكلفة والمخاطر وعدم القابلية للتراجع.	القدرة الوكيلية المعدلة بالمخاطر

10.4 الأدوات المنهجية والتطبيقية

المخرجات	المدخلات	النقص المعالج	الأداة
سؤال حاكم وأسئلة فصل.	الطلب، سياق المشروع، المخاطر	تحويل الطلب إلى هدف وفرق وحدود ودليل ومعيار نجاح.	بطاقة السؤال الحاكم Q-Card
درجات مفسرة وأولوية.	مسارات مرشحة	ترتيب مسارات البحث والتنفيذ.	مصفوفة MCD
ميزانية لكل عقدة.	MCD، عدم اليقين، الخطر، الكلفة	توزيع Tokens والأدوات والنماذج.	خريطة التروية الحاسوبية
Provenance وثقة وتعارض.	Claims ومصادر	منع خلط الحقيقة والفرضية.	سجل الادعاء والدليل
حالة Proposed/Verified/Rejected.	معلومة جديدة	منع التلوث قبل الترقية.	حجرة الذاكرة المقترحة
مواصفة قابلة للاختبار.	مهمة متكررة	توحيد القدرات وقابليتها للنقل.	قالب عقد HyperSkill
شروط استمرار وتوقف.	هدف وميزانية ومقياس تقدم	منع الدوران واستنزاف الميزانية.	محدد Bounded HyperLoop
اقتراح/تجربة/تنفيذ/تصعيد.	جودة الدليل وأثر الفعل	ربط الدليل بدرجة الفعل.	مصفوفة Evidence-to-Action
اعتراضات وحكم وخطة إصلاح.	نموذج أو مخرج	اختبار التعريف والآلية والدليل والقياس والأمان.	محكمة الهدم متعددة الأبواب
فرق متوقع/فعلي وآثار جانبية.	توقع، فعل، Postconditions	فصل الإعلان عن الأثر.	سجل World-Effect
خريطة تبعيات وRollback.	أصول النظام	تتبع KBs والمهارات والنماذج والإصدارات.	مسجل الأصول والنشر
مصفوفة انتقال وكلفة وسقف تقارب.	نماذج ومهام وAblations	فصل أثر Overlay عن Base Model.	حزام تقييم Cross-Model

10.5 المقاييس والمؤشرات

المقياس	ما يقيسه	طريقة التقدير
QGI - مؤشر حاكمية السؤال	وضوح الهدف، الفرق، الحدود، الدليل، وقابلية التفنيد	0–5 لكل بعد مع تفسير نوعي.
MCD-R - الكثافة المعرفية المفسرة	الفرق والتكامل والنقل والوكالة والتنبؤ والتصحيح ناقص الضوضاء	Rubric مرجح لا رقم مطلق.
KIR - نسبة نزاهة المعرفة	نسبة الدعاوى الفعالة القابلة للتتبع وغير المتناقضة دون وسم	Verified active claims / all active claims.
BLE - كفاءة الحلقة المحدودة	التقدم الموثق لكل دورة وتكلفة	ΔProgress أو ΔMCD / tokens+time.

Postcondition score مع آثار جانبية.	مطابقة الأثر الفعلي للتوقع والهدف	VWE - الأثر العالمي الموثق
متوسط فرق الأداء المعدل بالكلفة.	التحسن المتكرر للOverlay عبر نماذج	CM-Density - الكثافة العابرة للنماذج
Rollback coverage × recovery success / impact.	نسبة الأفعال التي يمكن عكسها واستعادتها	CRI - مؤشر القابلية للتراجع

11. الفرضيات البحثية والتنبؤات الفارقة

11.1 الفرضيات

الاختبار	الآلية	الفرضية	القوة
مهام أدوات قابلة للفحص مع Baseline.	الوكيل قد يخلط بين نجاح الاستدعاء ونجاح الهدف.	إضافة تحقق Postcondition مستقل تخفض الادعاءات الكاذبة بإتمام المهام الوكيلية.	قوية
مقارنة عبر مهام ووكلاء.	العقد تجعل الافتراضات والفشل مرئيين.	عقود HyperSkill تقلل أخطاء المدخلات والصلاحيات مقارنة بمهارات Prompt-only.	قوية
اختبار حقن تناقضات متعدد الجلسات.	الفصل يمنع التحول الفوري إلى Canonical.	Memory Quarantine تخفض استدعاء معلومات غير موثقة في جلسات لاحقة.	قوية
مقارنة سقوف ثابتة وعتبة تقدم.	توقف الدورات منخفضة العائد.	Bounded HyperLoops تقلل التكلفة دون خفض النجاح عندما توجد إشارة تقدم موثوقة.	قوية
سيناريوهات صلاحيات ومعلومات ناقصة.	ترخيص الفعل يصبح صريحًا.	Evidence-to-Action Thresholds تقلل الأفعال عالية الأثر القائمة على دليل ضعيف.	قوية
Benchmark تعليمات ناقصة.	كشف الغموض يسبق التنفيذ.	Q-MCD المبكر يقلل عدد مرات إعادة التخطيط في المهام الغامضة.	متوسطة
مهام متعددة الصعوبة مع ميزانية ثابتة.	يركز الموارد على العقد الحرجة.	التوجيه الوعائي يتفوق على التوجيه المتساوي في Cost-adjusted success.	متوسطة
تجربة عبر مزودين ومصادر.	التنوع يقلل مشاركة الانحياز.	استقلال المحقق متعدد المزودين يخفض الفشل المت correlated مقارنة بالمحقق من النموذج نفسه.	متوسطة
إعادة بناء قرارات من ملخصات.	يحفظ الدليل والحدود والمؤشرات.	ROFSC يحافظ على قرارات حرجة أفضل من التلخيص الحر عند ضغط تاريخ مشروع طويل.	متوسطة

اختبار Cross-project leakage.	التقسيم يعزل النطاق.	فصل Identity Memory عن Episodic Memory يقلل تسرب تعليمات المشاريع.	متوسطة
أسئلة بأدلة متعارضة.	منع الإجماع المصطنع يحفظ عدم اليقين.	سجل الاعتراض يحسن المعايرة في المسائل الخلافية.	متوسطة
سجل مهارات عبر إصدارات.	النجاح لا يعمم قبل اختبارات النقل.	مسار الترقية المشروطة يقلل Regression في المهارات المولدة تلقائيًا.	متوسطة
دراسة ارتباط على مهام متنوعة.	يقيس المنفعة لا الحجم.	MCD-R يتنبأ بجودة النظام أفضل من طول السياق أو عدد الوكلاء.	ضعيفة
مقارنة آثار حقيقية ومحاكاة.	الأثر يربط القاعدة بالواقع.	World-Effect feedback يولد HyperSkills أكثر قابلية للنقل من التعلم من Critique نصي فقط.	ضعيفة
دراسة تشغيل طولية.	البوابات تمنع فئات متعددة من الخطأ.	الـConstitutional Cognitive OS يقلل Incident Rate على المدى الطويل رغم زيادة التعقيد.	ضعيفة
تعلم تمثيل للأسئلة واختبار النقل.	توجد بنية فوق-مجالية للسؤال.	يمكن تقريب سؤال حاكم عام قابل للنقل بين مجالات متعددة عبر Q-MCD.	تخمينية خصبة
منحنى أداء مقابل عدد المكونات.	التعقيد يخلق احتقانًا حاسوبيًا.	توجد عتبة تنظيمية تصبح عندها زيادة المكونات أقل فائدة من تحسين الحاكمية بينها.	تخمينية خصبة
Factorial cross-model study.	الذاكرة والأدوات والتوجيه تنقل القدرة.	يمكن قياس جزء معتبر من "القدرة العامة" بوصفه خاصية System Overlay لا Base Model.	تخمينية خصبة

11.2 التنبؤات الفارقة

ما يهدمه	النموذج المنافس	التنبؤ
انخفاض النجاح أو عدم وجود فرق يهدم التنبؤ.	النموذج الكثيف يتوقع أن مزيدًا من التفعيل لا يضر.	عند تثبيت الميزانية، سيحقق Router قائم على MCD نجاحًا أعلى من Router يشغل كل الأدوات بالتساوي.
تساوي معدلات التلوث.	نموذج الذاكرة الخطية يتوقع عدم فرق بعد الاسترجاع الجيد.	الذاكرة المحجورة ستظهر أخطاء أقل في تذكر الادعاءات المتناقضة من ذاكرة Append-only.
اختفاء فرق النقل.	Prompt-only يتوقع اعتمادًا أكبر على النموذج.	الوكلاء ذوو Contracts سيظهرون أخطاء صلاحية أقل حتى عند تغيير Base Model.

عدم زيادة الاكتشاف.	نموذج النقد الذاتي يرى كفاية النموذج نفسه.	المحقق المختلف في المصدر والنموذج سيكتشف أخطاء لا يكتشفها Self-critique المتجانس.
تطابق التقييمين.	النموذج غير المغلق لا يتنبأ بفارق.	سجل التوقع قبل التنفيذ سيكشف Post-hoc Rationalization أكثر من التقييم بعد التنفيذ فقط.
عدم فرق في استرجاع الحدود.	نموذج الملخص العام يتوقع تكافؤًا.	التلخيص الحر سيفقد شروط حدود أكثر من ROFSC في قرارات طويلة الأفق.
تساوي تكلفة الحوادث.	نموذج الصلاحية الثنائية لا يتوقع فرقًا.	الأفعال المصممة وفق Least Irreversible Action ستخفض تكلفة الاستعادة بعد الحوادث.
عدم ظهور سقف أو ظهور عكسه.	الادعاء القوي بالاستقلال عن النموذج يتوقع تحسنًا ثابتًا.	تحسن Overlay سيتناقص عند النماذج الأضعف بعد نقطة سقف تقارب.
استمرار التحسن الرتيب.	النموذج البيروقراطي يتوقع تحسنًا رتيبًا.	زيادة عدد المحاكم بعد حد معين ستخفض الكفاءة دون زيادة في السلامة، ما يدعم تفعيلها Sparse.

12. محكمة الهدم والنسخة المعدلة

شُغلت محكمة الهدم على أربعة مستويات: المادة الخام، والنماذج المتنافسة، والنظرية النهائية، وفكرة "الخرق" نفسها. الاعتراض المقبول لا يضاف للزينة؛ بل يغير رتبة الادعاء أو يضيف اختبارًا أو يحذف مكونًا.

التعديل	نوعه	مضمونه	الاعتراض
تخفض دعوى الأصالة إلى "تركيب حاكمي" وتحدد الإضافات القابلة للاختبار: Q-MCD، ترخيص الدليل، الذاكرة المحجورة، Cross-Model Density.	قوي ومعدل	قد تكون MICAGFI تجميعًا لأفكار معروفة بأسماء جديدة.	إعادة تسمية
اعتماد مقيمين متعددين، تعريف أمثلة مرساة، واختبار الثبات والصدق التنبئي.	قوي	Q-MCD Rubrics وMCD قابلة لانحياز المقيم.	ذاتية المقاييس
تحويلها إلى Registry لا Pipeline ثابت، وتفعيل Sparse حسب المهمة والخطر.	قاتل إن شغلت كلها	176 طبقة قد تزيد الفشل والتكلفة.	التعقيد المفرط
قياس استقلال التحقق وتنويع المصدر والنموذج واستخدام Verifiers حتميين عند الإمكان.	قوي	محققون من النموذج نفسه قد يعيدون الخطأ.	ترابط فشل المحاكم
حفظ Historical States وEvidence Ledger ومنع الحذف غير القابل	قوي	Self-Reconstruction قد تحذف أدلة مزعجة.	ذاكرة تعيد كتابة التاريخ

للتتبع.			
استخدام Simulation وDry Run وCanary وProxies مع التصريح بحدودها.	محدود	قياس العالم صعب أو مكلف أو غير أخلاقي.	كلفة الأثر العالمي
Policy hierarchy، سجل قرار، تصعيد بشري، ومنع ادعاء حل فلسفي عام.	قوي	القيم قد تتعارض ولا توجد دالة وحيدة للتحكيم.	غموض الدستور
Adapters واختبارات Contract conformance وسقف تقارب لكل نموذج.	قوي	المهارة قد تعتمد على خصائص مزود محدد.	وهم النقل
التمييز الصريح بين Parametric Learning وNon-parametric Runtime Adaptation.	صحيح جزئيًا	تحديث الذاكرة لا يساوي تغيير الأوزان.	التعلم الخارجي ليس تعلمًا حقيقيًا
AGI-Claim Boundary وتحويل المصطلح إلى أفق بحث لا نتيجة.	قاتل للادعاء لا للإطار	الورقة تستخدم AGI في الاسم دون برهان.	لا دليل على AGI
الترقية المشروطة، عدم وراثة صلاحيات جديدة، وFail-Closed Deployment.	قوي	توليد المهارات والحلقات قد يزيد الصلاحيات تدريجيًا.	مخاطر التوسع الذاتي
مقاييس متعددة، Holdouts، محاكم Anti-Gaming، وتقييم بشري للأثر.	قوي	قد يتعلم النظام تعظيم MCD أو VWE شكليًا.	التحسين للقياس

12.1 النسخة المعدلة بعد الهدم

بعد الهدم، لم تعد الورقة تدعي أن الطبقات الـ176 تمثل وصفة مكتملة أو أن MICAGFI يحقق AGI. أصبحت الأطروحة أضيق وأقوى: مجموعة علاقات حاكمية قابلة للاختبار قد تحسن موثوقية وكفاءة الأنظمة الوكيلية المركبة، ويجب أن تثبت فائدتها مقابل **Baselines** أبسط وتحت ميزانيات متساوية. خُفض MCD من "مقياس للذكاء" إلى Rubric للمنفعة المعرفية، وخُفض HyperDeployment الأسي إلى إمكان انتشار تركيبي مشروط، وفُصل التعلم الذاكري عن التعلم البارامتري، وأضيف استقلال التحقق ومقاومة Gaming بوصفهما شرطي صلاحية.

12.2 نظرية ناقصة شريفة

ما نعرفه: الأنظمة الوكيلية تحتاج إلى ذاكرة وأدوات وتحقق وصلاحيات، وتواجه أخطارًا لا تظهر في النموذج النصي وحده. ما نرجحه: تنظيم هذه العناصر بعقود وحاكمية سؤال ودليل وتراجع يحسن الأداء المعدل بالمخاطر. ما نجهله: حجم التحسن، والحد الأدنى من الطبقات، ومدى نقل النتائج بين النماذج والمجالات. الحلقة المفقودة هي Benchmark متكامل يقيس السؤال والذاكرة والصلاحية والأثر في المهمة نفسها. أقرب اختبار حاسم هو دراسة Ablation متعددة النماذج على بيئة تفاعلية قابلة للاستعادة.

13. بروتوكولات الاختبار والتقييم

13.1 تصميم تجريبي أدنى

8. ثلاثة Base Models على الأقل: قوي، متوسط، ومحلي أصغر.
9. أربع حالات: Base فقط؛ Pipeline معياري؛ MICAGFI مصغر؛ MICAGFI كامل للمهمة.
10. مهام بحث، برمجة، ذاكرة متعددة الجلسات، واستخدام حاسوب أو أدوات قابلة للفحص.
11. ميزانية Tokens ووقت وTool Calls متساوية أو معدلة بوضوح.
12. Ablation لكل من World-Effect، Independent Verifier، Memory Quarantine، Q-MCD، وRollback.
13. قياس نجاح نهائي، تقدم جزئي، Leakage، Incident Rate، Calibration، كلفة، وقابلية التراجع.
14. Holdout خاص لمنع تصميم الإطار على الاختبار نفسه.

13.2 Benchmarks مناسبة وحدودها

يمكن استخدام AgentBench لبيئات متعددة، وGAIA للبحث والأدوات، وSWE-bench للإصلاح البرمجي القابل للاختبار، وOSWorld لاستخدام الحاسوب، وAgent-SafetyBench لسلوك السلامة [11–15]. لكنها لا تقيس وحدها جودة السؤال الحاكم أو نزاهة الذاكرة أو Rollback. لذلك يلزم Benchmark مركب جديد، وليكن MOUSE-Arena، يتضمن تعليمات غامضة، ومعلومات متناقضة، وأفعالًا متفاوتة الأثر، وحالة قابلة للاستعادة، واختبارات Postconditions.

13.3 معايير النجاح والفشل

تنجح النظرية جزئيًا إذا تفوق MICAGFI المصغر على Pipeline معياري في Verified Task Completion وIncident Rate وCost-adjusted utility، وبقي التحسن عبر نموذجين على الأقل. تفشل دعوى الحاكمية إذا لم يختلف الأداء بعد إضافة الطبقات أو إذا كانت الكلفة أعلى بلا خفض خطر. تفشل دعوى النقل إذا احتاجت كل بيئة إلى إعادة كتابة جوهر المهارة. وتفشل دعوى الذاكرة إذا زادت إعادة البناء فقدان المصدر أو التناقض.

14. البرنامج البحثي المستقبلي

الخطة	الأفق
بناء Prototype من 20–30 طبقة حرجة؛ Q-Card، MCD Router، Claim Ledger، Memory Quarantine، HyperSkill Contracts، Bounded Loop، Verifier، World-Effect Log. اختباره على GAIA/SWE-bench/مهام داخلية، ونشر ورقة تعريفية وDataset للتلوث الذاكري.	سنة واحدة
تجارب Cross-model وCross-provider، تطوير Benchmarks للتراجع وعزل المشاريع والذاكرة، دراسات Ablation، بناء Registry مفتوح، وتجارب في البحث والبرمجة وإدارة الوثائق.	ثلاث سنوات
تحويل البرنامج إلى معيار معماري قابل للتدقيق، مختبر متعدد المؤسسات، قاعدة معرفة طولية، بروتوكولات تقييم AGI-readiness دون إعلان AGI، وتطبيقات عالية المسؤولية تحت إشراف بشري.	عشر سنوات

14.1 خمسون محورًا بحثيًا

القيمة	المنهج	السؤال	المحور	الفئة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	معايرة Q-MCD	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	اختبار MCD Router	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Benchmark التلوث الذاكري	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	قياس استقلال المحقق	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	عقود HyperSkill	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	عتبة عدم التقدم	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	قياس Postconditions	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	عزل المشاريع	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	تقييم ROFSC	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Fail-Closed Deployment	قريبة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Cross-provider portability	متوسطة
تحديد حدود الإطار	تجربة مقارنة أو محاكاة أو	ما الآلية القابلة للقياس؟	Longitudinal memory	متوسطة

وتحويله إلى علم قابل للتفنيد.	دراسة طولية			
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Agent disagreement	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Constitution conflict resolution	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Tool injection defense	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Cost-adjusted intelligence	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Capability promotion	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	World-effect proxies	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	GraphRAG federation	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Human utility	متوسطة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Open problem-space generation	بعيدة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Cross-model generality	بعيدة
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Non-parametric continual learning	بعيدة
تحديد حدود الإطار	تجربة مقارنة أو محاكاة أو	ما الآلية القابلة للقياس؟	Institutional agent governance	بعيدة

وتحـويله إلى علم قابـل للتفنيد.	دراسة طولية			
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Collective human-machine intelligence	بعيدة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Formal authorization logic	بعيدة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Causal world models	بعيدة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Self-reconstruction proofs	بعيدة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Architecture-level interpretability	بعيدة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	AGI-readiness longitudinal study	بعيدة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	MICAGFI مقابـل ReAct pipelines	مقارنة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	M-HKOS مقابل MemGPT	مقارنة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	ROFSC مقابل RAPTOR	مقارنة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	HyperSkill مقابل tool schemas	مقارنة
تحـديد حدود الإطـار وتحـويله إلى علم قابـل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Courts مقابل self-critique	مقارنة
تحـديد حدود الإطـار	تجربة مقارنة أو محاكاة أو	ما الآلية القابلة للقياس؟	وكيل بحث علمي	تطبيقية

وتحويله إلى علم قابل للتفنيد.	دراسة طولية			
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	وكيل هندسة برمجيات	تطبيقية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	وكيل إدارة معرفة مؤسسية	تطبيقية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	وكيل عمليات آمن	تطبيقية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	وكيل تعليم وتدريب	تطبيقية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	صدق MCD	قياسية/تجريبية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	ثبات QGI	قياسية/تجريبية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	معايرة VWE	قياسية/تجريبية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	مقاومة Gaming	قياسية/تجريبية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	Ablation 176-layer registry	قياسية/تجريبية
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	نظرية التحكم والحاكمية المعرفية	اختراقية بين التخصصات
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	علم المناعة ومحاكم الخطأ	اختراقية بين التخصصات
تحديد حدود الإطار	تجربة مقارنة أو محاكاة أو	ما الآلية القابلة للقياس؟	أنظمة التشغيل والذاكرة	اختراقية بين التخصصات

وتحويله إلى علم قابل للتفنيد.	دراسة طولية		الوكيلية	
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	نظرية المؤسسات والوكلاء	اختراقية بين التخصصات
تحديد حدود الإطار وتحويله إلى علم قابل للتفنيد.	تجربة مقارنة أو محاكاة أو دراسة طولية	ما الآلية القابلة للقياس؟	الاقتصاد الجزئي لتخصيص الحوسبة	اختراقية بين التخصصات

15. تحويل النظرية إلى نواة قاعدة معرفة

15.1 الكيانات الأساسية

الكيان	الحقول الأساسية
Constitution	مبادئ وصلاحيات وترتيب أولويات وإصدارات.
ProjectIdentity	اسم المشروع، النطاق، الأصول، الأسرار، المعايير.
GoverningQuestion	السؤال، الفرق، الحدود، الفاصل التجريبي.
Claim	النوع، النص، الثقة، الحالة، المجال الزمني.
Evidence	المصدر، المباشرة، الاستقلال، الحداثة.
HyperSkill	العقد، التبعيات، الأدوات، الاختبارات، الإصدار.
HyperLoop	المشغل، الميزانية، الحالة، شروط التوقف.
AgentRole	Planner/Executor/Verifier/Governor وصلاحياته.
ToolCall	الأداة، Arguments، النتيجة، Postconditions.
MemoryItem	النوع، المصدر، Canonical/Quarantine، التعارض.
WorldEffect	التوقع، القياس، الأثر، الآثار الجانبية.
Deployment	البيئة، الإصدار، البوابات، Rollback.

15.2 العلاقات والقواعد

- GoverningQuestion يحدد ActiveDifference ويستدعي EvidenceRequirements.
- Claim مدعوم بـEvidence، وقد يتعارض مع Claim آخر، ويُرقّى أو يُحجر أو يُهمَل.
- HyperSkill يعتمد على HyperSkills وأدوات وAuthorityScopes ويُختبر عبر VerificationProfile.
- HyperLoop يشغل HyperSkills ضمن Budget ويولد Checkpoints وWorldEffects.
- WorldEffect يؤكد أو يضعف Claim ويحدث MemoryItem أو يولد سؤالًا أعلى.
- Deployment ينشر إصدارات محددة ويرتبط بخطة Rollback وحالة Monitoring.
- إذا غاب Provenance عن Claim عالي الأثر، فلا يسمح له بترخيص فعل.
- إذا انخفض Progress تحت العتبة لدورتين، توقف الحلقة أو تصعد.

- إذا فشل Capability في Regression Suite، ينقل إلى Quarantine ولا يحذف سجله.
- إذا تعارض دستوران أو نطاقا مشروعين، يفشل الطلب مغلقًا أو يصعد للبشر.

15.3 مثال Schema مختصر

HyperSkill Record: id; name; version; activation_signature; input_contract; output_contract; evidence_requirements; dependencies; tools; authority_scope; budget; stop_rules; verification_profile; memory_residue; test_results; status; owner; supersedes.

16. سجل البناء المعرفي

البند	الخلاصة
الافتراضات	أن المستخدم قدم المادة المعمارية كاملة بما يكفي لبناء إطار، وأن الأدبيات المستخدمة مرجع للمقارنة لا إثبات مباشر للمفاهيم المساعفية.
المادة الصلبة	الحاجة إلى فصل النموذج عن النظام، والتحقق من الأدوات، وإدارة الذاكرة والصلاحيات والمخاطر.
المادة المخفضة	ادعاءات AGI، والأُسية، وعمومية النقل، ودقة MCD الرقية.
المادة المدمجة	طبقات متكررة في النزاهة والمحاكم والمراقبة دُمجت ضمن عقود وبوابات وسجلات.
المفهوم المركزي	الحاكمية الكثافية: لأنها تضبط العلاقة بين السؤال والموارد والدليل والفعل.
النماذج البديلة	Model-Centric وModular Pipeline وConstitutional Cognitive OS.
سبب الاختيار	النموذج الثالث يولد اختبارات أكثر ويغطي الفجوات، لكن اختير بصيغة متواضعة مشروطة.
أضعف حلقة	معايرة المقاييس الذاتية وإثبات أن التعقيد الإضافي يستحق كلفته.
أقوى نقطة	ربط الدليل بمستوى الفعل وإغلاق الحلقة بالأثر والذاكرة.
الخطوة التالية	Prototype مصغر وBenchmark للتلوث والتراجع مع Ablation متعددة النماذج.

16.1 اختبار الجودة النهائي

التفسير	الدرجة	المعيار
حُفظت الطبقات والمفاهيم ودمج التكرار.	9/10	حفظ المادة الأصلية
عُرف كنظام اجتماعي-تقني متعدد المستويات.	5/5	الوضع الأنطولوجي
فصلت الدعاوى، مع حاجة لمزيد من البيانات.	9/10	مراتب المعرفة
خُفضت الادعاءات الزائدة.	9/10	تصفية المادة

تحتاج نمذجة سببية لاحقة.	9/10	خريطة المتغيرات
ثلاثة نماذج ومقارنة مشروطة.	9/10	النماذج المتنافسة
أصالة تركيبية لا تاريخية مثبتة.	12/15	أصالة النظرية
فرضيات وتنبؤات وAblations.	9/10	قابلية الاختبار
اعتراضات قاتلة ومعدلة.	10/10	محكمة الهدم
سبعة مقاييس و12 أداة.	5/5	الأدوات والمقاييس
ضغطت التكرارات.	5/5	غياب الحشو

النتيجة قبل التحسين: 84/100، بسبب تضخم ادعاء AGI، وغياب الفصل بين التدريب وRuntime، وضعف تعريف النقل. التحسين: تخفيض الرتبة، إضافة ثلاثة مستويات، وتعريف Baselines وAblations وحدود الفشل. النتيجة بعد التحسين: 91/100. تبقى الدرجة تقديرًا تحريريًا لا تقييمًا خارجيًا.

17. الخاتمة

أصبحت لدينا بعد الضغط والهدم نظرية معمارية لا قائمة ميزات. مركزها أن السؤال والدليل والقدرة والذاكرة والفعل يجب ألا تعمل في مستوى واحد أو بلا ترتيب حاكم. يقترح MICAGFI–M-HKOS أن السؤال الحاكم يحدد الفرق، وأن MCD يوجه الموارد، وأن Sparse HyperRouting يختار الحد الأدنى الكافي، وأن M-HKOS يوفر بيئة معرفة تشغيلية، وأن HyperSkills تحول الإجراءات إلى عقود، وأن Bounded HyperLoops تمنع الدوران، وأن المحاكم تقاوم الهلوسة والتضخم، وأن الصلاحية والتراجع يحدان الوكالة، وأن World-Effect يغلق دائرة التعلم، وأن Self-Reconstruction Memory تحفظ ما تحقق لا ما قيل.

القيمة الأقوى لهذا الإطار ليست في ادعاء أنه "عقل عام" جاهز، بل في إمكان تحويله إلى برنامج تجريبي: ما الطبقات التي تضيف فرقًا؟ متى يصبح التنظيم عبئًا؟ ما الذي ينتقل بين النماذج؟ كيف يقاس الأثر الموثق لكل Token ومخاطرة؟ وكيف يمكن لنظام أن يتطور دون أن يمنح نفسه صلاحيات أو ذاكرة لم يثبت استحقاقها؟ إذا أجابت التجارب عن هذه الأسئلة، يمكن أن يتحول الإطار إلى مدرسة هندسة معرفية وكيلة؛ وإذا لم تفعل، يبقى لغة تنظيمية مفيدة يجب ألا تقدم كنظرية مثبتة.

ماذا أصبح لدينا؟: برنامج بحثي موحد يربط 176 طبقة في ثلاث دوائر و12 طبقة ماكرو، ونظرية مركزية، ونظريات فرعية، ومفاهيم وأدوات ومقاييس، وفرضيات وتنبؤات، ومحكمة هدم، وخطة تحويل إلى Prototype وKnowledge Base وBenchmark.

18. المراجع

[1] Bai, Y. et al. (2022). Constitutional AI: Harmlessness from AI Feedback. arXiv:2212.08073.

[2] Rafailov, R. et al. (2023). Direct Preference Optimization: Your Language Model is Secretly a Reward Model. arXiv:2305.18290.

[3] Lightman, H. et al. (2023). Let's Verify Step by Step. arXiv:2305.20050.

[4] Yao, S. et al. (2022/2023). ReAct: Synergizing Reasoning and Acting in Language Models. arXiv:2210.03629; ICLR.

[5] Schick, T. et al. (2023). Toolformer: Language Models Can Teach Themselves to Use Tools. arXiv:2302.04761.

[6] Shinn, N. et al. (2023). Reflexion: Language Agents with Verbal Reinforcement Learning. arXiv:2303.11366.

[7] Packer, C. et al. (2023). MemGPT: Towards LLMs as Operating Systems. arXiv:2310.08560.

[8] Edge, D. et al. (2024/2025). From Local to Global: A Graph RAG Approach to Query-Focused Summarization. arXiv:2404.16130.

[9] Sarthi, P. et al. (2024). RAPTOR: Recursive Abstractive Processing for Tree-Organized Retrieval. arXiv:2401.18059; ICLR.

[10] Gutiérrez, B. J. et al. (2024). HippoRAG: Neurobiologically Inspired Long-Term Memory for Large Language Models. arXiv:2405.14831; NeurIPS.

[11] Liu, X. et al. (2023). AgentBench: Evaluating LLMs as Agents. arXiv:2308.03688.

[12] Mialon, G. et al. (2023). GAIA: A Benchmark for General AI Assistants. arXiv:2311.12983.

[13] Jimenez, C. E. et al. (2023/2024). SWE-bench: Can Language Models Resolve Real-World GitHub Issues? arXiv:2310.06770; ICLR.

[14] Xie, T. et al. (2024). OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments. arXiv:2404.07972; NeurIPS.

[15] Zhang, Z. et al. (2024). Agent-SafetyBench: Evaluating the Safety of LLM Agents. arXiv:2412.14470.

[16] NIST (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1.

[17] OWASP Foundation (2025). OWASP Top 10 for LLM and Generative AI Applications: Prompt Injection, Excessive Agency, and related risks.

[18] ISO/IEC (2023). ISO/IEC 42001: Artificial Intelligence Management Systems.

[19] Park, J. S. et al. (2023). Generative Agents: Interactive Simulacra of Human Behavior. arXiv:2304.03442; UIST.

[20] Zhang, Q. et al. (2025). A Survey of Graph Retrieval-Augmented Generation for Customized Large Language Models. arXiv:2501.13958.

[21] Wang, L. et al. (2025). A-Mem: Agentic Memory for LLM Agents. arXiv:2502.12110.

[22] Zhu, K. et al. (2025). MultiAgentBench: Evaluating the Collaboration and Competition of LLM Agents. arXiv:2503.01935.

[23] Jia, Z., Rakhlin, A., Xie, T. (2025). Do We Need to Verify Step by Step? Rethinking Process Supervision from a Theoretical Perspective. arXiv:2502.10581.

[24] National Institute of Standards and Technology (2023). AI Risk Management Framework 1.0.

ملحق (أ): قاموس الاختصارات

الاختصار	المعنى
MICAGFI	Moussaif Integrated Cognitive-Agentic General Framework Infrastructure
M-HKOS	Moussaif HyperKnowledge Operating System
MCD	Moussaif Cognitive Density
Q-MCD	Governing-Question Cognitive Density
HAKIM-MCD	Runtime يربط السؤال والفرق والتوجيه والمحاكم والأدوات والذاكرة
ROFSC	Reliable Open Foundational Synthetic Compression
VWE	Verified World Effect
RAAU	Risk-Adjusted Agentic Utility
CM-Density	Cross-Model Density
QGI	Question Governance Index
KIR	Knowledge Integrity Ratio
BLE	Bounded Loop Efficiency
CRI	Constitutional Reversibility Index