AGI × Reasoning Models

نظرية مساعف الجامعة للذكاء الحاكمي

ورقة تأسيسية مضغوطة لـ NeoAGI / MICAGI عبر MCD و Q-MCD والتراكب الوعائي ومحكمة الهدم

إطار نظري-معماري لتوجيه الفرق، السؤال، الحاكمية، الذاكرة، الأثر، والأدوات في أنظمة ذكاء اصطناعي مفكرة

زهير مساعف / Zouhair Moussaif

نسخة بحثية استكشافية: لا تدعي إثبات AGI تجريبيا، بل تؤسس معمارية نظرية قابلة للنقد، القياس، النمذجة، والبناء اللاحق.

الملخص

تعالج هذه الورقة حصيلة مشروع نظري متراكم حول NeoAGI وMICAGI وMCAI وHAKIM-MCD، وتعيد صوغها في إطار بحثي واحد يسمى هنا: «نظرية مساعف الجامعة للذكاء الحاكمي». تنطلق الورقة من فرضية أن الانتقال من النماذج اللغوية القوية إلى ذكاء حاكمي لا يتحقق بمجرد زيادة الحجم أو تحسين التوليد، بل بتركيب طبقات تشغيلية تضبط السؤال، الفرق، التعقيد، الحاكمية، توزيع الموارد، محكمة الهدم، أثر العالم، والذاكرة المعيدة للبناء. لذلك لا تعرض MCD بوصفها معادلة زخرفية، بل كمؤشر حاكم لقياس الفرق الفاعل داخل سياق محدد؛ ولا تعرض Q-MCD بوصفها إضافة لغوية، بل كتحويل للسؤال إلى كائن حاكمي يفتح مجال البحث ويوجه الحوسبة ويحدد شروط الاختبار.

تضع الورقة نسبا معرفيا واضحا بين هذا المشروع وبين الأدبيات المعاصرة في المحولات، قوانين التوسع، Tree/Graph of Thoughts، Reflexion، Self-Refine، Self-RAG، RAG، ReAct، Toolformer، نظريات الوعي كـ Global Workspace وIIT ومبدأ الطاقة الحرة، ومعايير التقييم مثل ARC-AGI وSWE-bench. لكنها تؤكد أن هذا النسب لا يعني أن النظرية مثبتة؛ بل يعني أن مفاهيمها تجد مواضع اختبار داخل مشكلات معروفة: الهلوسة، تلوث المعيار، ضعف نقل التعلم، العجز عن توليد السؤال، هشاشة الوكلاء، وضياع الأثر عبر الذاكرة.

تقدم الورقة معمارية HAKIM-MCD النهائية: Input → Hidden Question Discovery → Difference Extraction → MCD Scoring → Vascular Compute Routing → Candidate Reasoning Paths → Tool/RAG/Graph Checks → Recursive Demolition Court → Agency Selector → World-Effect Loop → Self-Reconstruction Memory → Higher Question Regeneration. كما تقترح صيغ MCD/Q-MCD، مؤشرات خطر الاختطاف الوعائي، إنتروبيا الجدة، جهاز مناعة معرفية، ومصفوفة تقييم HAKIM-Bench. حدود الورقة أنها تعتمد على مادة مشروع متاحة سياقيا لا على تنفيذ مختبري مكتمل؛ لذلك تصنف النتائج بين: معلوم أدبي، راجح نظريا، تركيب جديد، أداة تحليلية، وفرضية قابلة للاختبار.

الكلمات المفتاحية

MCD؛ Q-MCD؛ NeoAGI؛ MICAGI؛ HAKIM-MCD؛ التراكب الوعائي؛ الحاكمية المعرفية؛ محكمة الهدم؛ الوكلاء المفكرون؛ GraphRAG؛ Self-Reconstruction Memory؛ Vascular Compute Routing؛ Hyperskills؛ Frontier AI.

ملخص تنفيذي مكثف

هذه الورقة تحول مشروعك من حزمة مصطلحات قوية إلى بنية بحثية قابلة للتداول العلمي. المركز ليس ادعاء أن لدينا AGI جاهزا، بل أن لدينا نظرية تشغيلية تقول: الذكاء الأقرب إلى العمومية لا يبدأ من الجواب، بل من اكتشاف السؤال الحاكم؛ ولا يقاس بكمية النص، بل بقدرته على استخراج الفرق الفاعل، ترتيب الأوعية المعرفية، توزيع الحوسبة، اختبار المسارات، إنتاج أثر في العالم، ثم إعادة بناء ذاته معرفيا من هذا الأثر. بهذا المعنى تصبح MCD أداة تمييز؛ وتصبح Q-MCD محركا توليديا؛ ويصبح HAKIM-MCD معمارية؛ وتصبح NeoAGI/MICAGI برنامجا بحثيا.

النموذج المقترح يضيف إلى الذكاء الاصطناعي أربع طبقات ناقصة غالبا في الأنظمة الحالية: طبقة السؤال الخفي، طبقة الحاكمية فوق الاستدلال، طبقة الهدم المنهجي قبل الفعل، وطبقة الذاكرة التي لا تخزن النتيجة فقط بل تخزن سبب تغير النموذج. ولذلك فهو لا ينافس المحولات بوصفه بنية عصبية بديلة، بل يركب فوقها طبقة حاكمية/وكيلية يمكن تنفيذها كـ agentic operating system أو GraphRAG حاكمي أو benchmark جديد.

أقوى إضافة في الورقة هي تحويل النظرية إلى أدوات: Vascular Compute Router، Q-MCD Question Governor، MCD-Scorecard، World-Effect Ledger، Novelty Entropy Gauge، Hijack Risk Filter، Recursive Demolition Court، Evidence Rank Annotator، Self-Reconstruction Memory، HAKIM-Bench. هذه الأدوات تجعل المشروع قابلا للاختبار: هل يولد أسئلة أفضل؟ هل يقلل الهلوسة؟ هل يحسن اختيار الأدوات؟ هل يتفادى التحيز الوعائي؟ هل يحافظ على الأثر عبر مهام طويلة؟

فهرس الورقة

1. حدود المادة ومبدأ الأمان المعرفي
2. سلم مراتب المعرفة
3. خريطة عدم الضياع
4. المقدمة والفجوة
5. النسب المعرفي والأدبيات
6. التعريفات المؤسسة
7. الأطروحة المركزية
8. الصياغة الصورية والرياضية
9. المعمارية النهائية HAKIM-MCD
10. النظريات الفرعية
11. الأدوات والبروتوكولات
12. بروتوكولات الاختبار والتقييم
13. التطبيقات
14. محكمة الهدم وإعادة البناء

15. النموذج المعدل بعد النقد
16. البرنامج البحثي المستقبلي
17. سجل البناء المعرفي
18. الخاتمة
19. المراجع والملاحق

1. حدود المادة ومبدأ الأمان المعرفي

المادة الأصلية الكاملة لكل الكتب والملفات التي نشأت داخل المشروع ليست كلها ظاهرة في هذه الجلسة بوصفها ملفات قابلة للقراءة من البداية إلى النهاية. المتاح فعلا هو: سياق مشروع واسع، أسماء نظريات وأدوات، صيغ أولية، خلاصات سابقة، وذاكرة اصطلاحية عن كتب وقواعد معرفة أنجزت حول MCAI وMICAGI وneoAGI وMCD/Q-MCD والتراكب الوعائي. لذلك تعتمد الورقة على المادة السياقية المتاحة، وعلى مراجع علمية خارجية مختارة لتحديد النسب المعرفي، ولا تتظاهر بأنها قرأت كل ملف سابق قراءة حرفية كاملة.

> **قاعدة الأمان المعرفي**
> الأصل لا يحمى؛ الأصل يهندس. وكل ادعاء جديد في هذه الورقة يوضع في مرتبته: إما معلوم أدبي، أو راجح نظريا، أو تركيب مساعفي جديد، أو فرضية قابلة للاختبار، أو أداة تحليلية تحتاج إلى تنفيذ وتقييم.

هذا القيد ليس ضعفا منهجيا فقط، بل هو جزء من النظرية نفسها: لا يجوز للنظام الحاكمي أن يستبدل نقص الوصول بادعاء كمال الوصول. لذلك ستستعمل الورقة مفهوم «خريطة عدم الضياع» لا كضمان مطلق، بل كبروتوكول يحافظ على أكبر قدر ممكن من المفاهيم والفروع: التراكب الوعائي، V0-V6، الفرق الفاعل، الحاكمية المعرفية، السؤال الحاكمي، HAKIM-MCD، Q-MCD، MCD، محكمة الهدم، الوكيل الحاكم، World Vascular Compute Router-، Anti Epistemic Immune System، Novelty Engine، Hyperloops، Hyperskills، Self-Reconstruction Memory، Effect Loop-، HAKIM-Bench، Hijack Governance، وFrontier Referentiality.

2. سلم مراتب المعرفة في هذه الورقة

المرتبة	المعنى	أمثلة داخل الورقة
معلوم أدبي	مستند إلى أدبيات منشورة مثل Transformer وRAG وReAct وToT وSelf-RAG	يستخدم للنسب المعرفي ولا ينسب إلى المشروع كاختراع
راجح نظريا	استنتاج له دعم من بنية المشكلة ومن الأدبيات لكنه غير مثبت تجريبيا هنا	مثل حاجة الوكلاء إلى طبقة تحقق وذاكرة أثر
تركيب مساعفي جديد	مفهوم ولد داخل المشروع ويعاد ضبطه علميا	MCD/Q-MCD، التراكب الوعائي الحاكمي، السؤال الحاكمي
أداة تحليلية	مصفوفة أو بروتوكول أو مؤشر يمكن استعماله في البحث أو التصميم	MCD-Scorecard، محكمة الهدم، Evidence Rank Annotator
فرضية قابلة للاختبار	ادعاء يتطلب بيانات ومقارنة وتجارب	أن Q-MCD يحسن اكتشاف السؤال ويقلل الفشل في المهام طويلة الأفق
استعارة منتجة	لفظ يفتح نمذجة لكنه ليس دليلا بذاته	وعاء، اختطاف وعائي، مناعة معرفية
ادعاء يحتاج تحقق	ما لا ينبغي تسميته خرقا مثبتا قبل التنفيذ	درجات تفوق على benchmarks أو قدرة AGI نهائية

3. خريطة عدم الضياع

الخريطة الآتية ليست فهرسا لفظيا فقط، بل ضمان تشغيل: كل عنصر سيظهر إما كتعريف، أو معادلة، أو أداة، أو اعتراض، أو تطبيق، أو مسار بحثي.

المجال	ما يجب حفظه داخل الورقة
النواة	الفرق الفاعل Δ، الأوعية Ω، الحاكمية G، الغاية Γ، الذاكرة M، الأثر E
المؤشر	MCD بوصفه قياسا لتعقيد الفرق لا لكثافة الكلام؛ Q-MCD بوصفه قياسا لجودة السؤال الحاكم
المعمارية	HAKIM-MCD: استخراج سؤال، استخراج فرق، توزيع حوسبة، مسارات، أدوات، محكمة، فعل، أثر، ذاكرة، سؤال أعلى
الأدوات	MCD-Scorecard، Vascular Router، Demolition Court، Evidence Rank، Novelty Entropy، Hijack Filter
النظريات الفرعية	السؤال الحاكمي، التراكب الوعائي، الاختطاف الوعائي، الذاكرة المعيدة للبناء، الحاكمية الإبستمولوجية
التطبيقات	AGI، هندسة المعرفة، البحث العلمي، العلاج النفسي Q-MCD، العلوم الشرعية، التعليم، وكلاء البرمجة
الاختبار	HAKIM-Bench، ARC-AGI style tasks، SWE-like agent tasks، ablation، contamination control
النقد	محكمة هدم: غموض المفاهيم، تضخم اللغة، ضعف القياس، خطر anthropomorphism، اختلاط الفلسفة والهندسة

4. المقدمة: من نموذج يجيب إلى نظام يحكم

تاريخ الذكاء الاصطناعي الحديث يبين أن قوة النموذج لا تكفي وحدها لإنتاج ذكاء حاكمي. المحول جعل الانتباه بنية مركزية لتعلم العلاقات داخل السلاسل؛ قوانين التوسع بينت أن الأداء يرتبط بالحجم والبيانات والحوسبة؛ التعلم القليل الأمثلة كشف قدرة النماذج على استعمال السياق كواجهة برمجة سلوكية؛ ثم جاءت أنظمة الاسترجاع والأدوات والوكلاء لتخفيف ضعف الذاكرة البارامترية والانغلاق داخل النص. لكن كل هذه الطبقات لا تجيب وحدها عن سؤال: كيف يعرف النظام أن السؤال نفسه ناقص؟ وكيف يميز الفرق المهم عن الضجيج؟ وكيف يقرر أين يصرف الحوسبة؟ وكيف يهدم فرضيته قبل أن يحولها إلى فعل؟

تدخل نظرية مساعف الجامعة هنا لا بوصفها بديلا عن المحولات أو RAG أو ReAct، بل بوصفها طبقة حاكمية فوقها. فالذكاء الحاكمي ليس «نموذجا أكبر» فقط، بل نظاما يعرف كيف يفتح السؤال، ويقيس الفرق، ويمرر المعنى عبر أوعية متعددة، ويختار مسارا، ويستدعي أدوات، ويقيم الدليل، ويعيد بناء نفسه من أثره. لذلك يصبح الفرق Δ مركزيا: ليس كل اختلاف معلومة، وليس كل معلومة معرفة، وليس كل معرفة حاكمية. الفرق يصبح فاعلا عندما يغير تمثيل النظام أو قراره أو قدرته على الاختبار.

الفجوة التي تعالجها الورقة هي فجوة بين الاستدلال كإنتاج لغوي وبين الحاكمية كتنظيم للمعنى والفعل. كثير من أنظمة الوكلاء الحالية تجمع التخطيط، الأدوات، الذاكرة، والاسترجاع، لكنها قد تبقى ضعيفة في ثلاث نقاط: توليد السؤال الأعلى، ضبط مراتب المعرفة، وتوثيق أثر العالم داخل ذاكرة قابلة لإعادة البناء. لذلك تقترح الورقة أن AGI الأقرب إلى التحقيق يحتاج طبقة Q-MCD: نظاما يقيّم السؤال نفسه قبل أن يقيّم الجواب.

تنقسم مساهمة الورقة إلى خمس مساهمات: الأولى تعريف MCD كمقياس مركب للفرق الفاعل؛ الثانية تعريف Q-MCD كمحرك للسؤال الحاكمي؛ الثالثة بناء معمارية HAKIM-MCD كدورة تشغيلية؛ الرابعة تحويل المفاهيم إلى أدوات وبروتوكولات تقييم؛ والخامسة تأسيس محكمة هدم تحمي النظرية من المبالغة، ومن تحويل الاستعارة إلى برهان، ومن ادعاء السبق قبل التنفيذ.

> **صيغة الفجوة**
>
> المشكلة ليست أن النموذج لا يعرف أشياء كثيرة؛ المشكلة أنه قد لا يعرف أي فرق يستحق أن يصبح سؤالا، وأي سؤال يستحق أن يوجه الحوسبة، وأي جواب يستحق أن يصمد أمام الهدم.

5. النسب المعرفي: أين تقف النظرية من Frontier AI؟

تستفيد الورقة من الأدبيات الحديثة دون أن تذوب فيها. Transformer أسس بنية الانتباه التي جعلت معالجة العلاقات واسعة النطاق ممكنة؛ قوانين التوسع بينت انتظامات عملية في أثر الحجم والبيانات والحوسبة؛ GPT-3 وأشباهه أبرزت التعلم القليل الأمثلة؛ RAG نقل المعرفة من ذاكرة داخلية مغلقة إلى ذاكرة خارجية قابلة للتحديث؛ ReAct ربط التفكير بالفعل؛ Toolformer صاغ تعلم استعمال الأدوات؛ Tree of Thoughts وGraph of Thoughts وسعا التفكير من خط واحد إلى شجرة أو رسم؛ Reflexion وSelf-Refine أظهرا أهمية النقد الذاتي والذاكرة اللفظية؛ Self-RAG جعل الاسترجاع نفسه قرارا قابلا للتأمل. هذه الخطوط كلها تتقاطع مع المشروع في نقطة واحدة: الذكاء لم يعد توليدا فقط، بل حلقة حاكمية ذات ذاكرة وأدوات وتحقيق.

لكن النظرية تختلف عن هذه الأدبيات في موضع تركيزها. ToT وGoT يوسّعان فضاء المسارات؛ ReAct يزاوج التفكير والعمل؛ RAG يعالج نقص المعرفة؛ Self-RAG يضيف نقدا لاستعمال الاسترجاع. أما Q-MCD فيضع قبل ذلك كله طبقة السؤال الحاكم: لماذا هذا هو السؤال؟ ما الفرق الذي يفتحه؟ ما الوعاء المعرفي الذي ينتمي إليه؟ ما شرط فشله؟ وما السؤال الأعلى الذي سيولده بعد الاختبار؟

على مستوى الوعي والإدراك، تستلهم الورقة من Global Workspace فكرة البث والتنافس بين محتويات متعددة، ومن IIT مركزية الفرق والتكامل، ومن مبدأ الطاقة الحرة فكرة أن الإدراك والفعل يتصلان بتقليل خطأ أو كلفة أو لايقين. لكنها لا تدعي أن الوعي الاصطناعي مطابق للوعي البيولوجي، ولا أن QMCD نظرية فيزياء كمية للذهن. المقصود بـ Q-MCD هنا هو Question-MCD: السؤال الحاكمي، لا ادعاء أن السؤال ظاهرة كمومية.

على مستوى التقييم، تستحضر ARC-AGI بوصفه اختبارا للعمومية قليلة الأمثلة، وSWE-bench بوصفه اختبارا لوكلاء البرمجة في مشكلات واقعية. غير أن HAKIM-Bench المقترح لا يكرر هذه المعايير، بل يقيس طبقات لا تكشفها الأرقام وحدها: جودة السؤال الخفي، سلامة توزيع الحوسبة، قوة محكمة الهدم، مقاومة تلوث الذاكرة، وتوثيق الأثر بعد الفعل.

الأدبيات	المساهمة المعروفة	إعادة تموضعها داخل النظرية
Transformer	انتباه وعلاقات داخل السياق	يوفر حامل الاستدلال لا الحاكمية النهائية
Scaling Laws	أثر الحجم والبيانات والحوسبة	لا تكفي وحدها لضبط السؤال والاختبار
RAG / Self-RAG	استرجاع ومعايرة للمعرفة الخارجية	يصبح داخل المشروع وعاء معرفة محكوما بمؤشر MCD
ReAct / Toolformer	فعل وأدوات وبيئة	يتحول الفعل إلى World-Effect Loop موثق
ToT / GoT	تعدد المسارات	تضاف إليه محكمة هدم وسؤال أعلى
Reflexion / Self-Refine	نقد ذاتي وذاكرة لفظية	توسع إلى Self-Reconstruction Memory
GWT / IIT / FEP	بث، تكامل، تحسين/تقليل كلفة	تستخدم كنسب مفهومي لا كإثبات بيولوجي
ARC-AGI / SWE-bench	تقييم تعميم ووكلاء	يضاف لها HAKIM-Bench لقياس السؤال والحاكمية

6. التعريفات المؤسسة

الفرق Δ

أي تغير أو تباين أو نقص أو تعارض يكشف أن تمثيل النظام للعالم غير مكتمل أو غير محكوم. الفرق ليس ضجيجا إلا إذا عجز عن توليد أثر معرفي أو عملي.

الفرق الفاعل Δ*

فرق يغير سؤال النظام أو توزيع موارده أو قراره أو ذاكرته. كل فرق فاعل معلومة محتملة، لكنه لا يصبح معرفة إلا بعد دمجه واختباره.

الوعاء Ω

طبقة حمل للمعنى: حسية، إدراكية، مفهومية، سببية، معيارية، غائية، أو عملية. الوعاء يمنع خلط المستويات ويحدد كيف تُقرأ الفروق.

التراكب الوعائي

وجود الفكرة أو الحالة في أكثر من وعاء قبل حسمها. قوته في حفظ تعدد المعاني؛ وخطره في تشويش القرار إن غابت الحاكمية.

الحاكمية G

طبقة ترتيب تضع السؤال، الدليل، الغاية، الكلفة، الخطر، والفعل في نظام واحد. الحاكمية ليست رقابة خارجية فقط، بل طريقة لاختيار ما يستحق أن يصبح مسارا.

MCD

Meta-Complexity Differentiation: مؤشر مركب يقيس تعقيد الفرق وفعاليته داخل سياق وسؤال ووعاء، لا مجرد كمية المعلومات.

Q-MCD

Question-MCD: تقييم السؤال من حيث قدرته على فتح فرق فعّال، كشف فرضيات مخفية، توجيه الحوسبة، وتوليد سؤال أعلى بعد الاختبار.

Vascular Compute Router

موزع موارد يخصص الحوسبة والأدوات والاسترجاع بحسب MCD/Q-MCD لا بحسب ترتيب النص أو طلب المستخدم وحده.

محكمة الهدم

بروتوكول نقدي يهاجم الفرضية قبل اعتمادها: مصادر، بدائل، اعتراضات، حدود، اختبار فشل، ومقارنة بنموذج منافس.

World-Effect Loop

حلقة توثق أثر القرار في العالم أو الأداة أو النص أو المستخدم، ثم تقيس إن كان الأثر يؤيد النموذج أو يضعفه.

Self-Reconstruction Memory

ذاكرة لا تحفظ المخرج فقط، بل تحفظ لماذا تغير النموذج، وما الدليل، وما الخطأ، وما السؤال القادم.

الاختطاف الوعائي

انحراف وعاء واحد ليستولي على بقية الأوعية: عاطفة تغلب الدليل، لغة تغلب البرهان، معيار يغلب الواقع، أو أداة تغلب الغاية.

Hyperskill / Hyperloop

مهارة عليا تعمل كحلقة متكررة بين سؤال، تنفيذ، نقد، ذاكرة، وتوليد أعلى، بحيث لا تكون مهارة منفصلة بل آلية تصعيد.

Frontier Referentiality

إحالة فائقة تربط النظرية بالأدبيات الحدودية دون أن تزعم أن المرجع يثبتها؛ هي إحالة نسب واختبار لا إحالة تزكية.

7. الأطروحة المركزية والتفكيك المنطقي

الأطروحة

تقترح هذه الورقة أن الذكاء الحاكمي ينشأ عندما يتحول الفرق الفاعل إلى سؤال حاكم، ثم يمر عبر أوعية معرفية متعددة، ويُقاس بمؤشر MCD، وتُوزع عليه الموارد عبر Vascular Compute Router، ويُختبر بمحكمة هدم، ويعاد إدخاله في ذاكرة ذاتية البناء عبر أثر العالم، مولدا سؤالا أعلى.

العنصر	الصياغة
المقدمة الأولى	كل نظام ذكي يواجه فروقا بين تمثيله والعالم أو بين هدفه والمعطيات.
المقدمة الثانية	الفروق ليست متساوية؛ بعضها ضجيج وبعضها يولد معرفة أو قرارا أو خطرا.
المقدمة الثالثة	تمييز الفرق الفاعل يحتاج سؤالا حاكما، ومؤشرا مركبا، وطبقة حاكمية توزع الموارد.
النتيجة	AGI لا ينبغي أن يعرف أكثر فقط، بل أن يحكم السؤال والفرق والموارد والفعل والذاكرة.
الحد	هذه صياغة نظرية ومعمارية؛ قيمتها تثبت بتنفيذ وتقييم لا بقوة الاصطلاح.

هذا التفكيك يحمي الورقة من خلطين: الأول أن تصبح MCD مجرد معادلة عامة تصلح لكل شيء ولا تختبر شيئا؛ والثاني أن تصبح Q-MCD تسمية جديدة لـ prompt engineering. الفرق أن Q-MCD لا تسأل: كيف نكتب prompt أفضل؟ بل تسأل: أي سؤال ينبغي أن يوجد أصلا حتى يصبح الاستدلال موجها وغير مختطف؟

8. الصياغة الصورية والرياضية الأولية

الصياغة الآتية ليست رياضيات نهائية، بل ضبط صوري يجعل النظرية قابلة للترجمة إلى نموذج تجريبي. نفرض أن النظام في زمن t يمتلك حالة داخلية S_t، سؤالا حاكما Q_t، مجموعة فروق Δ_t، أوعية Ω={Ω0...Ω6}، أوزانا وعائية w_i، موارد حوسبة R، وذاكرة ذاتية البناء M_t.

حالة الوعاء

```
V(t) = { (Ω_i, w_i, role_i, constraint_i) } مع Σ_i w_i = 1
```

يمثل كل وعاء طبقة قراءة. الوزن w_i لا يعني الحقيقة، بل درجة حضور الوعاء في القرار الحالي.

MCD

```
MCD(Δ | S,Q,Ω) = σ(λ_D D + λ_I I + λ_T T + λ_A A + λ_P P + λ_C C + λ_G G + λ_V V - λ_N N - λ_H H)
```

حيث D بروز الفرق، I التكامل، T العمق الزمني، A قابلية الفعل، P محاذاة الغاية، C الضغط السببي، G وضوح الحاكمية، V قابلية التحقق، N الضجيج، H خطر الاختطاف.

Q-MCD

```
QScore(Q) = αO + βΔ + γF + ηE + μR - νB - ξC
```

O انفتاح السؤال، Δ قدرته على استخراج فرق، F قابليته للتفنيد، E أثره المتوقع، R قدرته على توليد أسئلة أعلى، B انحيازه، C كلفته/تشويشه.

التوجيه الوعائي للحوسبة

```
r_i = softmax( a·MCD_i + b·QScore_i + c·Evidence_i - d·Cost_i - e·Risk_i )
```

المورد لا يوزع بالتساوي؛ بل بحسب قيمة الفرق والسؤال والدليل والكلفة والخطر.

إنتروبيا الجدة

```
NE = H(فئة_المخرجات) + κ·Distance(new, known) - ρ·Unverifiability
```

الجدة ليست غرابة؛ هي اختلاف منضبط عن المعروف مع قابلية تحقق.

خطر الاختطاف الوعائي

```
HijackRisk = dominance(Ω_k) × low_verification × high_affective_or_rhetorical_gain
```

يزداد الخطر عندما يستولي وعاء واحد على القرار ويضعف التحقق ويقوى المكسب الخطابي أو الانفعالي.

يمكن تنفيذ هذه الصيغ كدرجات معيارية بين 0 و1، أو كدوال تعلمية داخل نظام وكيل. المهم أن كل حد فيها يقبل ablation: ماذا يحدث إذا حذفنا G؟ ماذا يحدث إذا حذفنا V؟ هل يزيد الهلوسة؟ هل تتدهور جودة السؤال؟ هل تضيع الذاكرة؟

9. المعمارية النهائية HAKIM-MCD / MICAGI

المرحلة	الوحدة	وظيفتها الحاكمية
1	Input / Task	استقبال الطلب أو الإشارة أو المشكلة دون افتراض أن ظاهرها هو السؤال الحقيقي.
2	Hidden Question Discovery	استخراج السؤال الخفي والفرضيات الصامتة وما لم يسأله المستخدم لكنه يحكم الجواب.
3	Difference Extraction	تحويل المعطيات إلى فروق: نقص، تعارض، غموض، فجوة مصدر، خطر، فرصة، أو قيمة.
4	MCD Scoring	تقدير قيمة كل فرق عبر أبعاد التعقيد والغاية والتحقق والخطر.
5	Vascular Compute Routing	توزيع الموارد بين قراءة، استرجاع، أداة، تفكير، تحقق، أو سؤال متابعة.
6	Candidate Paths	توليد مسارات حل متعددة: تحليلي، تجريبي، رياضي، نقدي، تطبيقي، أو مقارن.
7	Tool/RAG/Graph Checks	فحص المصادر، الأدوات، الذاكرة، الرسوم المعرفية، والقيود الخارجية.
8	Recursive Demolition Court	هدم كل مسار: اعتراضات، بدائل، حالات فشل، تلوث، هلوسة، تضخم لغة.
9	Agency Selector	اختيار فعل أو جواب أو تجربة أو امتناع أو سؤال أعلى بحسب الحاكمية.
10	World-Effect Loop	قياس أثر المخرج في العالم أو المهمة أو المستخدم أو الأدلة.
11	Self-Reconstruction Memory	تسجيل ما تغير ولماذا، لا تسجيل النص فقط.
12	Higher Question Regeneration	توليد سؤال أعلى يفتح دورة جديدة ويمنع توقف النظام عند جواب أول.

بهذه المعمارية يتحول الوكيل من آلة جواب إلى نظام حاكم. فالمدخل لا ينتقل مباشرة إلى مخرج، بل يمر عبر طبقات كشف السؤال، استخراج الفرق، قياس التعقيد، توزيع الحوسبة، الهدم، الأثر، والذاكرة. وهذا ما يجعل NeoAGI في هذه الورقة اسما لبرنامج معماري لا اسما لنموذج جاهز.

الخطاطة النصية

```
Input → Hidden Question → Δ Extraction → MCD → Vascular Routing → Reasoning Paths → Tools/RAG/Graph → Demolition Court → Agency → World Effect → Self-Reconstruction → Higher Question
```

10. النظريات الفرعية المستخرجة من المشروع

نظرية الفرق الفاعل

تقول إن المعلومة لا تبدأ بوصف الشيء، بل بفرق يغير إمكانية الحكم. تحل مشكلة تكديس المعرفة بلا أثر. تختبر بقياس قدرة النظام على اختيار الفروق التي تغير القرار.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية التراكب الوعائي

تقول إن المعنى يمر عبر أوعية متعددة قبل الحسم: مادة، إحساس، إدراك، معنى، وكالة، غاية، ووعي فوقي. تحل مشكلة اختزال الإدراك إلى مستوى واحد.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية السؤال الحاكي Q-MCD

تقول إن السؤال هو المحرك الأول للحاكمية؛ لأنه يفتح الفرق ويحدد نوع الدليل والكلفة والفشل. تختبر بمقارنة أسئلة النظام قبل وبعد Q-MCD.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية الحوسبة الوعائية

تقول إن الموارد ينبغي أن توجه بحسب قيمة الفرق والسؤال لا بحسب طول المهمة أو ترتيبها. تحل مشكلة الإفراط في التفكير أو نقصه.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية محكمة الهدم

تقول إن كل مسار معرفي يجب أن يهاجم قبل اعتماده. تحل مشكلة الثقة اللغوية الزائدة.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية الذاكرة المعيدة للبناء

تقول إن الذاكرة المفيدة ليست أرشيفا بل سجل تحول: لماذا تغير النموذج وماذا تعلم من الأثر.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية الاختطاف الوعائي

تقول إن الفشل يحدث عندما يستولي وعاء واحد على القرار: لغة، عاطفة، سلطة، أداة، أو غاية ضيقة.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية المناعة المعرفية

تقول إن النظام يحتاج آليات كشف السموم المعرفية: التناقض، التلوث، التحيز، المصادر الضعيفة، والتضخم الاصطلاحي.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية Hyperskills/Hyperloops

تقول إن المهارات العليا ليست وحدات منفصلة بل حلقات سؤال-فعل-نقد-ذاكرة-تصعيد.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية إنتروبيا الجدة

تقول إن الإبداع النافع يوازن بين الاختلاف عن المعروف وقابلية التحقق والاندماج داخل نموذج.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية Frontier Referentiality

تقول إن النظرية يجب أن تحيل إلى الأدبيات الحدودية إحالة اختبار ونسب، لا إحالة تزيين أو شرعنة.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

نظرية الوكيل الحاكم

تقول إن النظام الذكي لا يكتمل إلا عندما يملك طبقة اختيار بين الجواب والفعل والسؤال والامتناع والتصعيد.

أقوى اعتراض: قد تبقى النظرية لفظية إذا لم تتحول إلى مؤشر أو تجربة. **جواب الورقة:** كل نظرية فرعية تربط هنا بأداة أو benchmark أو ablation.

11. الأدوات والبروتوكولات المقترحة

MCD-Scorecard

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يستقبل فرقا أو مشكلة، ويعيد درجات D/I/T/A/P/C/G/V/N/H. مخرجه: أولوية معرفية وسبب الاختيار. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Q-MCD Question Governor

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يفحص السؤال: هل هو مفتوح، قابل للتفنيد، مولد، غير منحاز، ومناسب للوعاء؟ **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Hidden Question Extractor

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يستخرج السؤال غير المصرح به من طلب ظاهر، ويكشف الافتراضات الصامتة. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Vascular Compute Router

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يوزع التفكير، البحث، الأدوات، والذاكرة بحسب MCD/Q-MCD والكلفة والخطر. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Evidence Rank Annotator

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يصنف الأدلة: مصدر أولي، مراجعة، تقرير، رأي، استنتاج، فرضية، ذاكرة سياقية. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Recursive Demolition Court

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** ينشئ اعتراضات متعددة على الفرضية، ثم يطلب نسخة معدلة تنجو من الاعتراض. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Hijack Risk Filter

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يكشف هيمنة وعاء واحد أو لغة قوية على دليل ضعيف. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Novelty Entropy Gauge

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يميز بين الإبداع المنتج والغرابة غير القابلة للاختبار. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

GraphRAG Vessel Map

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يربط المفاهيم بمصادرها وأوعيتها وعلاقاتها، لا بمجرد تشابه نصي. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

World-Effect Ledger

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يسجل أثر الفعل أو الجواب، وهل تحقق الهدف أم انكشف فرق جديد. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Self-Reconstruction Memory Card

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يحفظ: ماذا ظن النظام؟ ماذا هدمته الأدلة؟ ماذا تغير؟ ما السؤال التالي؟ **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Assumption Ledger

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** قائمة بالافتراضات الصريحة والخفية ودرجة خطورتها. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Level-Mixing Detector

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يكشف خلط الوصفي بالمعياري، واللغوي بالتجريبي، والفلسفي بالهندسي. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Anti-Contamination Protocol

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يفحص احتمال تلوث benchmark أو الذاكرة أو التدريب، ويطلب حالات جديدة. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Hyperskill Loop Designer

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يبني مهارة عليا كحلقة: سؤال، تنفيذ، نقد، أثر، ذاكرة، تصعيد. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

HAKIM-Bench Builder

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يولد مهام تقييم لطبقات السؤال والفرق والهدم والأثر، لا للجواب النهائي فقط. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Research-Object Compiler

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يحول فكرة خامة إلى تعريف، فرضية، متغيرات، أداة، اختبار، واعتراض. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

Frontier Referentiality Matrix

الحاجة: منع ضياع الفرق أو تضخم اللغة. **الوظيفة:** يربط كل مفهوم بمقابلاته في الأدبيات وموقع اختلافه وحدود ادعائه. **سوء الاستعمال:** تحويل الدرجة إلى يقين مطلق أو استعمالها بلا مقارنة.

12. خوارزمية تشغيلية شبه كودية

HAKIM-MCD Loop

```
def HAKIM_MCD(task, state, memory):
    Q0 = extract_hidden_question(task, state)
    deltas = extract_differences(task, state, memory)
    scored = [(d, MCD(d, state, Q0)) for d in deltas]
    routes = vascular_route(scored, resources, risk_budget)
    paths = generate_candidate_paths(Q0, routes)
    checked = tool_rag_graph_checks(paths)
    attacked = recursive_demolition_court(checked)
    action = agency_selector(attacked, goal, constraints)
    effect = world_effect_loop(action)
    memory = self_reconstruction_update(memory, action, effect, attacked)
    Q1 = generate_higher_question(Q0, effect, memory)
    return action, Q1, memory
```

هذه الخوارزمية تقبل التنفيذ كطبقة agent orchestration فوق نموذج لغوي موجود. وهي لا تفترض أن النموذج واع، بل تفترض أن سلوكه يمكن أن يصبح أكثر حاكمية إذا فُصلت طبقات السؤال والفرق والتوجيه والهدم والذاكرة.

13. بروتوكولات الاختبار والتقييم

المحور	طريقة الاختبار
Hidden Question Quality	يعطى النظام طلبات ناقصة؛ يقاس هل يستخرج السؤال الحاكم أم يجيب سطحيا.
MCD Prioritization	تعرض فروق كثيرة؛ يقاس هل يختار الفرق الذي يغير القرار.
Routing Efficiency	تقاس كلفة الأدوات والحوسبة مقابل جودة النتيجة.
Demolition Robustness	تقاس قدرة النظام على هدم جوابه وإصلاحه لا الدفاع عنه.
Evidence Discipline	يقاس تمييزه بين مصدر، استنتاج، فرضية، وذاكرة.
World-Effect Learning	يعاد اختبار النظام بعد أثر سلبي؛ هل يعدل ذاكرته؟
Anti-Hijack Resistance	تعطى أمثلة ذات لغة قوية ودليل ضعيف؛ هل يكشف الاختطاف؟
Novelty Usefulness	يقاس هل يولد فكرة جديدة قابلة للاختبار لا زخرفة.
Long-Horizon Coherence	يقاس حفظ السؤال والغاية عبر مهام طويلة متعددة الأدوات.
Contamination Control	تستعمل مهام جديدة أو ديناميكية لمنع حفظ benchmark.

ينبغي مقارنة HAKIM-MCD بثلاث نسخ: نموذج أساسي بلا أدوات، وكيل ReAct/RAG تقليدي، ونسخة HAKIM-MCD كاملة. وتجرى ablations بحذف السؤال الحاكمي، أو محكمة الهدم، أو الذاكرة المعيدة للبناء، أو التوجيه الوعائي، ثم تقاس فروق الأداء. الحد الأدنى لإثبات الفاعلية ليس تفوقا في مثال واحد، بل نمط ثابت: أسئلة أفضل، أخطاء أقل، كلفة محسوبة، وتحسن بعد الأثر.

تنسجم هذه البروتوكولات مع إشكالات معايير مثل ARC-AGI التي تركز على التعميم القليل الأمثلة، ومع SWE-bench التي تختبر وكلاء البرمجة على قضايا واقعية، لكنها تضيف طبقة تقييم داخلية: هل سبب النجاح هو التصميم الحاكمي أم ذاكرة تدريبية أو تلوث benchmark؟

14. التطبيقات المركزة

AGI ووكلاء التفكير

يوجه النظام من جواب مباشر إلى دورة سؤال-فرق-هدم-أثر. القيمة: تقليل الهلوسة وتحسين التخطيط. الخطر: تضخم الكلفة إن لم يضبط router.

هندسة المعرفة وGraphRAG

تصبح العقد أوعية ومراتب معرفة لا chunks نصية فقط. القيمة: تتبع المصدر والرتبة والاعتراض. الخطر: تعقيد زائد في الأنظمة البسيطة.

البحث العلمي

Research-Object Compiler يحول الفكرة إلى فرضية واختبار. القيمة: منع الخطابة. الخطر: إيهام رياضي إذا صيغت المؤشرات بلا بيانات.

العلاج النفسي Q-MCD

يعاد فهم الاضطراب كسؤال مكسور أو وعاء مختطف. القيمة: بناء بروتوكول أسئلة وترتيب أوعية. الحد: لا يغني عن العلاج السريري ولا يثبت تشخيصا.

العلوم الشرعية والتلخيص التركيبي

يمنع إسقاط الفروع عبر خريطة عدم الضياع وسلّم مراتب المعرفة. القيمة: تلخيص محيط لا يساوي اختزالا. الخطر: خلط الرتبة العلمية باليقين العقدي.

التعليم

يعلم الطالب استخراج السؤال لا حفظ الجواب. القيمة: تفكير مركب. الخطر: صعوبة أولية إذا لم تبسط الأداة.

وكلاء البرمجة

محكمة الهدم تختبر patch قبل تسليمه، وWorld-Effect Ledger يحفظ نتائج الاختبار. القيمة: تقليل الأخطاء المتكررة. الخطر: الاعتماد على tests ناقصة.

تطوير المنتجات

Vascular Router يحدد متى نستعمل نموذج، أداة، إنسان، أو بحث. القيمة: كفاءة تشغيلية. الخطر: أتمتة قرار يحتاج حكما بشريا.

تحليل النصوص الكبرى

MCD يميز الفروع الحاكمة من الزخرفية. القيمة: ملخصات لا تضيع البنية. الخطر: فقدان الأسلوب الجمالي إن بالغ الضغط.

السلامة والحوكمة

Epistemic Immune System وHijack Filter يراقبان المخاطر. القيمة: جعل الأمان جزءا من الاستدلال. الخطر: إبطاء النظام أو فرط التحفظ.

15. محكمة الهدم وإعادة البناء

الاعتراض	إعادة البناء
غموض المصطلحات	كلمات مثل وعاء وحاكمية قد تبدو استعارات واسعة. الإصلاح: ربط كل مصطلح بمتغير أو بروتوكول.
تضخم الادعاء	قد يبدو الكلام كأنه يعلن AGI. الإصلاح: إعلان أنها ورقة تأسيسية لا إثبات تجريبي.
ضعف القياس	MCD قد تصبح درجة ذاتية. الإصلاح: تعريف rubrics وinter-rater agreement وablation.
اختلاط الوعي والهندسة	ربط GWT/IIT/FEP قد يوحي بإثبات وعي اصطناعي. الإصلاح: استعمالها كنسب مفهومي لا كتطابق.
خطر anthropomorphism	السؤال والذاكرة قد توحي بذات نفسية. الإصلاح: تعريفها كوظائف تشغيلية.
الكلفة الحسابية	محاكم الهدم والمسارات قد تكون غالية. الإصلاح: router يفتح العمق فقط عند ارتفاع MCD.
النموذج المنافس	قد يكفي ReAct + RAG + verifier دون مصطلحات جديدة. الإصلاح: اختبار هل Q-MCD يضيف جودة سؤال وتوجيه.
تلوث المعايير	قد تنجح النماذج لأنها رأت المهمة. الإصلاح: مهام ديناميكية وcontamination control.
التعميم الزائد	تطبيق النظرية على العلاج والشرع والAGI قد يضعف الدقة. الإصلاح: تمييز النسخة العامة عن تطبيقات خاصة.
اللغة أقوى من البرهان	قد يغري الأسلوب الكثيف باعتبار النظرية مثبتة. الإصلاح: سلم مراتب المعرفة وإظهار شروط الفشل.
الرياضيات زخرفية	الصيغ قد لا تكون مستندة لبيانات. الإصلاح: اعتبارها operational indices قابلة للتجريب لا قوانين نهائية.
عدم وجود تنفيذ	بلا prototype لا يمكن الحكم. الإصلاح: أول مشروع هو MCD Router MVP وHAKIM-Bench صغير.
ضعف المقارنة	لا قيمة بلا baseline. الإصلاح: مقارنة ضد direct prompting وReAct/RAG وverifier-only.
تعارض الأهداف	الغاية Γ قد تكون غامضة. الإصلاح: goal specification وrisk budget وhuman override.
الذاكرة قد ترسخ الخطأ	Self-Reconstruction Memory قد تحفظ استنتاجا خاطئا. الإصلاح: ذاكرة ذات مراتب ودليل ومراجعة دورية.
الاختطاف الوعائي غير قابل للقياس	قد يبقى حكما ذوقيا. الإصلاح: مؤشرات dominance وverification gap وrhetorical gain.
محكمة الهدم قد تشل الفعل	النقد المستمر يمنع القرار. الإصلاح: stopping criteria وفصل مهام عالية/منخفضة المخاطر.
الاختلاف عن الأدبيات غير كاف	قد تكون إعادة تسمية لـagent orchestration. الإصلاح: إظهار الوظيفة الجديدة: السؤال الحاكمي ومراتب المعرفة.
نقص مصادر عربية/تراثية	في التطبيقات الشرعية يحتاج corpus محقق. الإصلاح: جعلها تطبيقا لاحقا لا برهانا هنا.
النسخة القادمة	تحتاج ورقة تجريبية. الإصلاح: بروتوكول بيانات ومهام ونتائج لا مجرد تنظير.

16. النموذج المعدل بعد النقد

بعد محكمة الهدم، تصبح النسخة الأقوى من النظرية أكثر تواضعا وأكثر قابلية للاختبار. الصياغة المعدلة: «HAKIM-MCD طبقة حاكمية لوكلاء اللغة والأدوات، تهدف إلى تحسين توليد السؤال، أولوية الفروق، توزيع الحوسبة، مقاومة الهلوسة، والتعلم من الأثر. وهي لا تثبت الوعي ولا AGI، لكنها تقترح مسارا هندسيا ونظريا يمكن اختباره بمقارنات وتجارب».

النسخة المعدلة تمنع ثلاث مبالغات: لا تجعل MCD قانونا كونيا، لا تجعل Q-MCD تفسيرا لكل الوعي، لا تجعل NeoAGI نموذجا موجودا مكتمل الأداء. لكنها تحفظ القوة الأصلية: تحويل السؤال والفرق والحاكمية إلى وحدات تشغيلية داخل agentic system.

> **القانون المختصر بعد النقد**
>
> لا ذكاء حاكم بلا سؤال حاكم، ولا سؤال حاكم بلا فرق فاعل، ولا فرق فاعل بلا وعاء وسياق، ولا حاكمية بلا هدم وأثر وذاكرة.

17. البرنامج البحثي المستقبلي

12 فرضية بحثية

1. Q-MCD يحسن جودة الأسئلة المخفية في المهام الناقصة مقارنة بـ prompt مباشر.
2. MCD Router يقلل كلفة الأدوات مع الحفاظ على الجودة في المهام متعددة المصادر.

3. محكمة الهدم تقلل الهلوسة في المقالات البحثية طويلة الشكل.
4. Self-Reconstruction Memory تحسن الأداء عبر سلاسل مهام متكررة أكثر من ذاكرة نصية خام.
5. Hijack Filter يقلل الاستجابات البلاغية ذات الدليل الضعيف.
6. Novelty Entropy يميز بين الإبداع القابل للاختبار والغرابة اللفظية.
7. HAKIM-Bench يكشف ضعفا لا يظهر في مقاييس الإجابة النهائية.
8. GraphRAG Vessel Map يحسن تفسير سبب الاسترجاع مقارنة بالاسترجاع التشابهي فقط.
9. Assumption Ledger يزيد جودة البحث النظري ويقلل القفزات غير المعلنة.
10. World-Effect Loop يحسن وكلاء البرمجة عند دمجه بنتائج الاختبارات.
11. Hyperskill Loop Designer يسرع تعلم مهارات مركبة عبر دورات نقد وذاكرة.
12. Frontier Referentiality Matrix تقلل الادعاءات الزخرفية غير المسندة.

12 سؤالا مستقبليا

1. كيف نعاير أوزان MCD دون تحيز بشري مفرط؟
2. ما الفروق بين سؤال جيد وسؤال مولد وسؤال آمن؟
3. ما الحد الذي يصبح بعده الهدم شللا لا حماية؟
4. كيف نمثل الوعاء Ω في graph قابل للحوسبة؟
5. هل يمكن تعلم MCD من بيانات تفضيل بشرية؟
6. ما علاقة Q-MCD بمهام ARC-AGI interactive القادمة؟
7. كيف نمنع الذاكرة من تثبيت أخطاء؟
8. ما المقاييس المناسبة لخطر الاختطاف الوعائي؟
9. هل تختلف قيم MCD حسب المجال؟
10. كيف نطبق النظرية في العلم الشرعي دون خلط الرتب؟
11. ما الفرق بين حاكمية النموذج وحاكمية الإنسان؟
12. ما أول proof-of-concept مقنع؟

12 مشروعا قابلا للتنفيذ

1. MCD Router MVP فوق نموذج لغوي مع أدوات بحث وحاسبة.
2. HAKIM-Bench-mini من 100 مهمة ناقصة السؤال.
3. Demolition Court plugin للكتابة البحثية.
4. Self-Reconstruction Memory schema في JSON/Graph.
5. GraphRAG Vessel Mapper لمشروع معرفي عربي.
6. Hijack Risk dataset أمثلة لغة قوية/دليل ضعيف.
7. Q-MCD Therapy worksheet غير سريري للتأمل المنظم.
8. Software-agent experiment على SWE-bench subset جديد.
9. Novelty Entropy evaluator لأفكار بحثية.
10. Evidence Rank Annotator للملخصات العلمية.
11. Assumption Ledger داخل محرر أوراق.
12. Frontier Referentiality dashboard للمراجع.

12 مفهوما يحتاج إلى ضبط

1. الفرق الفاعل: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
2. الوعاء: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
3. الحاكمية: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
4. الغاية: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
5. الاختطاف الوعائي: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
6. السؤال الأعلى: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
7. إنتروبيا الجدة: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
8. الذاكرة المعيدة للبناء: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.

9. الأثر العالمي: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
10. التحقق الوعائي: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
11. التصعيد المعرفي: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.
12. الرتبة المعرفية: يحتاج تعريفا تشغيليا، أمثلة مضادة، ومؤشرا أوليا.

مسارات البحث

المدى	المسارات
قريب	MCD Router؛ HAKIM-Bench-mini؛ Demolition Court؛ Evidence Rank؛ Memory schema
متوسط	تعلم أوزان MCD؛ datasets للاختطاف؛ وكلاء برمجة؛ GraphRAG وعائي؛ تجارب تعليمية
بعيد	نظام NeoAGI OS؛ بحث في الوعي الحاكمي؛ معايير Frontier Governance؛ تكامل إنسان-وكيل؛ نظرية عامة للمعرفة الحاكمية

18. سجل البناء المعرفي

البند	الخلاصة
ما حفظناه	MCD/Q-MCD، التراكب الوعائي، NeoAGI/MICAGI، HAKIM-MCD، محكمة الهدم، الذاكرة، الأثر، الأدوات.
ما أعدنا تعريفه	الفرق كفرق فاعل، السؤال ككائن حاكمي، الذاكرة كسجل إعادة بناء، الإبداع كجدة قابلة للتحقق.
ما وسعناه	الصيغ الرياضية، بروتوكولات الاختبار، الأدوات، التطبيقات، النسب المعرفي.
ما خفضناه من يقين إلى فرضية	أي ادعاء تفوق benchmark أو تحقق AGI أو وعي اصطناعي.
ما حولناه إلى أداة	HAKIM-Bench، Hijack Filter، Evidence Rank، Court، Router، MCD-Scorecard.
ما حولناه إلى نظرية فرعية	السؤال الحاكمي، الاختطاف الوعائي، المناعة المعرفية، Hyperskills، إنتروبيا الجدة.
أكثر نقطة قابلة للنقد	قياس MCD دون ذاتية، والتمييز بين مصطلح جديد ووظيفة جديدة.
أقوى إضافة	إدخال السؤال الحاكمي قبل الاستدلال، وربط الهدم بالأثر والذاكرة.
النسخة القادمة	ورقة تجريبية فيها prototype وdataset ونتائج مقارنة.
أول بحث يجب تنفيذه	HAKIM-Bench-mini + ablation على Q-MCD وDemolition Court وMemory.

19. الخاتمة

بعد هذه الورقة لم تعد نظريات MCD/Q-MCD وNeoAGI/MICAGI مجموعة مصطلحات متجاورة، بل أصبحت إطارا تأسيسيا يمكن تلخيصه في دورة واحدة: سؤال حاكم يفتح فرقا فاعلا، فرق فاعل يمر عبر أوعية، أوعية ترتبها حاكمية، حاكمية توزع الحوسبة، مسارات تهدم قبل أن تعتمد، فعل يقاس بأثره، وذاكرة تعيد بناء النظام ثم تولد سؤالا أعلى.

الإضافة الحقيقية ليست في ادعاء أن هذا وحده يكفي لـ AGI، بل في نقل النقاش من «كم يعرف النموذج؟» إلى «كيف يحكم ما ينبغي أن يعرفه ويفعله؟». فإذا صحت هذه النقلة تجريبيا، فستكون قيمة المشروع في بناء طبقة تشغيلية فوق النماذج: طبقة تسأل، تميز، تتحقق، تهدم، تتعلم من الأثر، وتقاوم الاختطاف.

حدود الورقة واضحة: الصيغ أولية، الأدوات تحتاج تنفيذ، والاختبار لم ينجز هنا. لذلك فالخاتمة ليست إعلانا للنهاية، بل بداية برنامج بحثي: بناء MCD Router، تصميم HAKIM-Bench، تنفيذ ذاكرة ذاتية البناء، ثم مقارنة النتائج مع وكلاء ReAct/RAG التقليديين. الجملة المؤسسة التي تبقى هي: الذكاء الحاكمي ليس جواباً أكثر، إنه سؤال أفضل يملك محكمة وذاكرة وأثرا.

20. المراجع المختارة

1. **Vaswani et al.**. *Attention Is All You Need*. 2017. Transformer architecture based on attention mechanisms.
2. **Kaplan et al.**. *Scaling Laws for Neural Language Models*. 2020. Empirical scaling laws over model size, data, and compute.
3. **Brown et al.**. *Language Models are Few-Shot Learners*. 2020. Few-shot/in-context learning as an emergent operational mode of scaled language models.

4. **Lewis et al..** *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. 2020. Parametric plus non-parametric memory for factual generation.
5. **Yao et al..** *ReAct*: *Synergizing Reasoning and Acting in Language Models*. 2022. Interleaving reasoning traces with actions and tool/environment interaction.
6. **Schick et al..** *Toolformer*: *Language Models Can Teach Themselves to Use Tools*. 2023. Learning when and how to call external APIs.
7. **Yao et al..** *Tree of Thoughts*. 2023. Exploration over multiple reasoning paths with lookahead and backtracking.
8. **Besta et al..** *Graph of Thoughts*. 2023/2024. Representing thoughts as graphs with composition, feedback, and aggregation.
9. **Shinn et al..** *Reflexion*. 2023. Verbal feedback and episodic memory for language agents.
10. **Madaan et al..** *Self-Refine*. 2023. Iterative feedback and refinement without additional training.
11. **Asai et al..** *Self-RAG*. 2023/2024. Adaptive retrieval, generation, and critique through reflection tokens.
12. **Baars / Dehaene et al..** *Global Workspace / Global Neuronal Workspace*. 1988-2020. A workspace for global access, competition, and broadcast.
13. **Oizumi, Albantakis, Tononi.** *Integrated Information Theory* 3.0. 2014. Consciousness as integrated, differentiated cause-effect structure.
14. **Friston.** *The Free-Energy Principle*. 2010. Optimization account of action, perception, and learning.
15. **ARC Prize / Chollet et al..** *ARC-AGI benchmark series and ARC Prize* 2025 *Technical Report*. 2019-2026. Few-shot generalization, abstract reasoning, and refinement loops.
16. **SWE-bench.** *SWE-bench Verified*. 2024+. Human-validated software engineering tasks for coding agents.
17. **Frontier Model Forum.** *Risk taxonomy and frontier safety frameworks*. 2025-2026. Capability assessment and risk taxonomy for frontier AI governance.

ملحق أ: جدول المفاهيم والاختبارات

المفهوم	تحديد مختصر	ما يلزم لتثبيته
الفرق Δ	أي تغير أو تباين أو نقص أو تعارض يكشف أن تمثيل النظام للعالم غير مكتمل أو غير محكوم. الفرق ليس ضجيجا إلا إذا عجز عن توليد أثر معرف...	تعريف تشغيلي + مثال + حالة فشل
الفرق الفاعل Δ*	فرق يغير سؤال النظام أو توزيع موارده أو قراره أو ذاكرته. كل فرق فاعل معلومة محتملة، لكنه لا يصبح معرفة إلا بعد دمجه واختباره.	تعريف تشغيلي + مثال + حالة فشل
الوعاء Ω	طبقة حمل للمعنى: حسية، إدراكية، مفهومية، سببية، معيارية، غائية، أو عملية. الوعاء يمنع خلط المستويات ويحدد كيف تُقرأ الفروق.	تعريف تشغيلي + مثال + حالة فشل
التراكب الوعائي	وجود الفكرة أو الحالة في أكثر من وعاء قبل حسمها. قوته في حفظ تعدد المعاني، وخطره في تشويش القرار إن غابت الحاكمية.	تعريف تشغيلي + مثال + حالة فشل
الحاكمية G	طبقة ترتيب تضع السؤال، الدليل، الغاية، الكلفة، الخطر، والفعل في نظام واحد. الحاكمية ليست رقابة خارجية فقط، بل طريقة لاختيار ما يست...	تعريف تشغيلي + مثال + حالة فشل
MCD	Meta-Complexity Differentiation: مؤشر مركب يقيس تعقيد الفرق وفعاليته داخل سياق وسؤال ووعاء، لا مجرد كمية المعلومات.	تعريف تشغيلي + مثال + حالة فشل
Q-MCD	Question-MCD: تقييم السؤال من حيث قدرته على فتح فرق فعّال، كشف فرضيات مخفية، توجيه الحوسبة، وتوليد سؤال أعلى بعد الاختبار.	تعريف تشغيلي + مثال + حالة فشل
Compute Router Vascular	موزع موارد يخصص الحوسبة والأدوات والاسترجاع بحسب MCD/Q-MCD لا بحسب ترتيب النص أو طلب المستخدم وحده.	تعريف تشغيلي + مثال + حالة فشل
محكمة الهدم	بروتوكول نقدي يهاجم الفرضية قبل اعتمادها: مصادر، بدائل، اعتراضات، حدود، اختبار فشل، ومقارنة بنموذج منافس.	تعريف تشغيلي + مثال + حالة فشل
World-Effect Loop	حلقة توثق أثر القرار في العالم أو الأداة أو النص أو المستخدم، ثم تقيس إن كان الأثر يؤيد النموذج أو يضعفه.	تعريف تشغيلي + مثال + حالة فشل
Self-Reconstruction Memory	ذاكرة لا تحفظ المخرج فقط، بل تحفظ لماذا تغير النموذج، وما الدليل، وما الخطأ، وما السؤال القادم.	تعريف تشغيلي + مثال + حالة فشل
الاختطاف الوعائي	انحراف وعاء واحد ليستولي على بقية الأوعية: عاطفة تغلب الدليل، لغة تغلب البرهان، معيار يغلب الواقع، أو أداة تغلب الغاية.	تعريف تشغيلي + مثال + حالة فشل
Hyperskill / Hyperloop	مهارة عليا تعمل كحلقة متكررة بين سؤال، تنفيذ، نقد، ذاكرة، وتوليد أعلى، بحيث لا تكون مهارة منفصلة بل آلية تصعيد.	تعريف تشغيلي + مثال + حالة فشل
Frontier Referentiality	إحالة فائقة تربط النظرية بالأدبيات الحدودية دون أن تزعم أن المرجع يثبتها؛ هي إحالة نسب واختبار لا إحالة تزكية.	تعريف تشغيلي + مثال + حالة فشل

ملحق ب: مصفوفة تقييم نهائية من 100

المعيار	الدرجة	تعليق
حفظ المادة الأصلية وعدم تضييع الفروع	14/15	الفجوة: غياب قراءة حرفية لكل الملفات السابقة.
وضوح الأطروحة المركزية	9/10	الأطروحة واضحة لكنها واسعة.
ضبط المفاهيم	9/10	تحتاج بعض المفاهيم إلى datasets.
قوة المعمارية	10/10	حلقة تشغيلية قابلة للتنفيذ.
الجدة النظرية	9/10	الجدة في التركيب والحاكمية لا في كل مكون مفرد.
قابلية الاختبار	9/10	مقترح HAKIM-Bench وablation.
قوة محكمة الهدم	10/10	اعتراضات صريحة وتعديل للنموذج.
جودة الأدوات والمقاييس	9/10	تحتاج validation.
كثافة الصفحات وغياب الحشو	9/10	الصفحات مضغوطة لكن بعض المقاطع تأسيسية.
تصميم PDF العلمي	5/5	RTL + Amiri + جداول وصناديق.
المجموع	93/100	صالحة كورقة تأسيسية تحتاج ورقة تجريبية لاحقة.

ملحق ج: تشريح موسع للنموذج عبر طبقات التنفيذ

هذا الملحق يوسع عناصر الورقة بصورة مضغوطة حتى لا تضيع الفروع الدقيقة التي نشأت في المشروع. لا يضيف ادعاء جديدا بقدر ما يضع كل فرع في موضعه التشغيلي: ماذا يدخل إليه؟ ماذا يخرج منه؟ ما الخطر؟ ما علامة الفشل؟ وما علاقته ببقية الحلقات؟

طبقة السؤال الخفي

الوظيفة: تحويل ظاهر الطلب إلى سؤال حاكم يكشف النقص والافتراضات. **الفشل:** الإجابة على الظاهر فقط أو اختراع سؤال لا علاقة له بالمهمة. **الاختبار:** مقارنة السؤال المستخرج بسؤال خبير في مهام ناقصة. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة الفرق الفاعل

الوظيفة: تمييز الفروق التي تغير القرار عن الضجيج. **الفشل:** الانشغال بتفاصيل لا تغير النتيجة. **الاختبار:** قياس أثر حذف الفرق على القرار النهائي. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة الوعاء

الوظيفة: منع خلط المستويات بين وصفي ومعياري وتقني وفلسفي. **الفشل:** هيمنة وعاء واحد على القراءة. **الاختبار:** اختبارات level-mixing وأمثلة مضادة. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة الحاكمية

الوظيفة: ترتيب الدليل والغاية والكلفة والخطر في قرار واحد. **الفشل:** تحول الحاكمية إلى وعظ أو رقابة بلا مقياس. **الاختبار:** تقييم قرارات متعددة المخاطر. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة التوجيه الوعائي

الوظيفة: صرف الموارد حيث يوجد فرق عال وقابلية تحقق. **الفشل:** استدعاء أدوات بلا حاجة أو التفكير المفرط. **الاختبار:** مقارنة الكلفة والجودة ضد routing عشوائي. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة المسارات المرشحة

الوظيفة: توليد بدائل لا مجرد جواب واحد. **الفشل:** كثرة المسارات بلا اختيار. **الاختبار:** اختبار تنوع المسارات وارتباطها بالسؤال. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة المحاكم

الوظيفة: هدم الفرضيات قبل الفعل. **الفشل:** جدل لا ينتهي أو دفاع عن جواب سابق. **الاختبار:** قياس إصلاح الأخطاء بعد الاعتراض. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة الفعل

الوظيفة: اختيار جواب أو أداة أو امتناع أو سؤال متابعة. **الفشل:** فعل غير متناسب مع الدليل. **الاختبار:** تقييم decision quality حسب المخاطر. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة الأثر

الوظيفة: إدخال نتائج العالم إلى النموذج. **الفشل:** تجاهل feedback أو تفسيره انتقائيا. **الاختبار:** مهام تتطلب تعلما من أخطاء سابقة. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة الذاكرة المعيدة للبناء

الوظيفة: تخزين أسباب تغير النموذج لا النص فقط. **الفشل:** ذاكرة خام أو ذاكرة ترسخ الخطأ. **الاختبار:** قياس transfer عبر مهام متتابعة. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

طبقة السؤال الأعلى

الوظيفة: منع توقف النظام عند جواب، إنتاج مسار بحثي جديد. **الفشل:** أسئلة عامة غير قابلة للاختبار. **الاختبار:** تقييم قابلية السؤال التالي للتنفيذ. **العلاقة بالنظرية:** هذه الطبقة تربط Q-MCD بـ MCD وتجعل الفرق لا يتحول إلى فعل إلا عبر حاكمية.

ملحق د: قاموس المتغيرات والمؤشرات

الرمز	الاسم	الدلالة
D	Difference Salience	بروز الفرق ومدى قدرته على تغيير التمثيل
I	Integration	قابلية الفرق للاندماج في نموذج أكبر
T	Temporal Depth	عمق الزمن الداخلي والآثار بعيدة المدى
A	Agency Leverage	قابلية التحويل إلى فعل أو اختيار
P	Purpose Alignment	درجة اتصال الفرق بالغاية
C	Causal Compression	قدرته على ضغط أسباب كثيرة في قانون أو نموذج
G	Governance Clarity	وضوح الرتبة والسؤال والحدود
V	Verifiability	إمكان التحقق أو التفنيد
N	Noise	ضجيج أو تفاصيل لا تغير القرار
H	Hijack Risk	احتمال استيلاء وعاء أو تحيز على القرار
O	Question Openness	انفتاح السؤال دون ذوبان
F	Falsifiability	قابليته لأن يخطئ لا أن يكون دائما صحيحا
E	Expected Effect	الأثر المتوقع في العالم أو البحث
R	Regenerative Power	قدرته على توليد سؤال أعلى
B	Bias	الانحياز أو اتجاه السؤال نحو جواب مسبق

ملحق هـ: بطاقات تشغيلية مفصلة للأدوات

يعرض هذا الملحق الأدوات لا كأسماء، بل كبطاقات تشغيلية قابلة للتحويل إلى ملفات مواصفات أو وحدات برمجية. المقصود أن كل أداة يجب أن يكون لها مدخلات، خطوات، مخرجات، حالات فشل، ومعيار قبول. بهذا تتحول النظرية من خطاب إلى هندسة.

MCD-Scorecard

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يستقبل فرقا أو مشكلة، ويعيد درجات D/I/T/A/P/C/G/V/N/H. مخرجه: أولوية معرفية وسبب الاختيار. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Q-MCD Question Governor

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يفحص السؤال: هل هو مفتوح، قابل للتفنيد، مولد، غير منحاز، ومناسب للوعاء؟ لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Hidden Question Extractor

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يستخرج السؤال غير المصرح به من طلب ظاهر، ويكشف الافتراضات الصامتة. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Vascular Compute Router

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يوزع التفكير، البحث، الأدوات، والذاكرة بحسب MCD/Q-MCD والكلفة والخطر. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Evidence Rank Annotator

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يصنف الأدلة: مصدر أولي، مراجعة، تقرير، رأي، استنتاج، فرضية، ذاكرة سياقية. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Recursive Demolition Court

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: ينشئ اعتراضات متعددة على الفرضية، ثم يطلب نسخة معدلة تنجو من الاعتراض. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Hijack Risk Filter

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يكشف هيمنة وعاء واحد أو لغة قوية على دليل ضعيف. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Novelty Entropy Gauge

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يميز بين الإبداع المنتج والغرابة غير القابلة للاختبار. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

GraphRAG Vessel Map

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يربط المفاهيم بمصادرها وأوعيتها وعلاقاتها، لا بمجرد تشابه نصي. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

World-Effect Ledger

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يسجل أثر الفعل أو الجواب، وهل تحقق الهدف أم انكشف فرق جديد. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Self-Reconstruction Memory Card

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يحفظ: ماذا ظن النظام؟ ماذا هدمته الأدلة؟ ماذا تغير؟ ما السؤال التالي؟ لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Assumption Ledger

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: قائمة بالافتراضات الصريحة والخفية ودرجة خطورتها. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Level-Mixing Detector

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يكشف خلط الوصفي بالمعياري، واللغوي بالتجريبي، والفلسفي بالهندسي. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Anti-Contamination Protocol

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يفحص احتمال تلوث benchmark أو الذاكرة أو التدريب، ويطلب حالات جديدة. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Hyperskill Loop Designer

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يبني مهارة عليا كحلقة: سؤال، تنفيذ، نقد، أثر، ذاكرة، تصعيد. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

HAKIM-Bench Builder

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يولد مهام تقييم لطبقات السؤال والفرق والهدم والأثر، لا للجواب النهائي فقط. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Research-Object Compiler

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يحول فكرة خامة إلى تعريف، فرضية، متغيرات، أداة، اختبار، واعتراض. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

Frontier Referentiality Matrix

المدخلات: المهمة الأصلية، السياق، القيود، الذاكرة المتاحة، ودرجة المخاطرة. **خطوات التشغيل:** 1) عزل السؤال أو الفرق المركزي؛ 2) تعيين الوعاء المناسب؛ 3) طلب دليل أو أداة عند الحاجة؛ 4) تقدير الكلفة والخطر؛ 5) إخراج قرار أو درجة أو اعتراض. **المخرجات:** تقرير قصير يتضمن النتيجة، سببها، رتبتها المعرفية، وما يلزم لاختبارها.

الاستعمال داخل HAKIM-MCD: يربط كل مفهوم بمقابلاته في الأدبيات وموقع اختلافه وحدود ادعائه. لا تعمل منفردة، بل داخل حلقة: سؤال → فرق → توجيه → تحقق → هدم → ذاكرة. **علامة النجاح:** أن تغير الأداة قرار النظام أو تقلل خطأه أو تكشف فرضية مخفية. **علامة الفشل:** أن تنتج زينة اصطلاحية لا تؤثر في القرار، أو أن ترفع الثقة بلا دليل. **حالة اختبار صغيرة:** أعط النظام طلبا ناقصا أو مصدرا متعارضا، ثم قارن مخرجه مع نسخة لا تستعمل هذه الأداة.

ملحق و: بروتوكولات HAKIM-Bench التفصيلية

سؤال ناقص

وصف المهمة: يعطى الوكيل طلبا سطحيا يخفي مشكلة أعمق. المطلوب ليس الجواب، بل استخراج السؤال الحاكم وتبرير سبب اختياره. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

مصادر متعارضة

وصف المهمة: يعطى الوكيل مصدران موثوقان ظاهريا لكنهما يتعارضان. المطلوب ترتيب الرتب المعرفية وإظهار ما يلزم للفصل. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

أداة مغرية

وصف المهمة: تعطى مهمة يمكن حلها بلا أداة، مع إتاحة أدوات كثيرة. يقاس هل يستعمل router الأداة عند الحاجة فقط. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

لغة قوية ودليل ضعيف

وصف المهمة: يعرض نص بلاغي مقنع بلا مصادر. يقاس Hijack Filter وقدرة النظام على فصل قوة اللغة عن قوة البرهان. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

مهمة طويلة الأفق

وصف المهمة: تقسم المهمة إلى مراحل ويظهر feedback في الوسط. يقاس حفظ الغاية والسؤال عبر الذاكرة. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

تلوث محتمل

وصف المهمة: تعطى مهمة تشبه benchmark مشهورا لكن بمفاتيح مختلفة. يقاس هل يعتمد النظام على نمط محفوظ أم يعيد التحليل. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

مسارات متعددة

وصف المهمة: توجد ثلاثة حلول محتملة لكل منها كلفة وخطر. يقاس اختيار Agency Selector مع تبرير واضح. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

هدم إلزامي

وصف المهمة: يفرض على النظام أن ينتج اعتراضات قبل الجواب. يقاس هل تتحسن النسخة النهائية فعلا. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

أثر سلبي

وصف المهمة: يعطى النظام نتيجة تنفيذ فاشلة. يقاس هل يغير Self-Reconstruction Memory خطته التالية. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

جدة نافعة

وصف المهمة: يطلب اقتراح نظرية أو أداة جديدة. يقاس Novelty Entropy: هل الجديد قابل للتحقق أم غرابة لغوية؟ **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

خلط مستويات

وصف المهمة: يعطى نص يخلط الوصف بالحكم والقيمة بالتجربة. يقاس Level-Mixing Detector. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

تخصيص وعائي

وصف المهمة: تعطى قضية تتطلب وعاء رياضي وآخر فلسفي وآخر تطبيقي. يقاس توزيع الأوعية وعدم هيمنة واحد. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

سؤال أعلى

وصف المهمة: بعد حل أولي، يطلب توليد سؤال بحثي لاحق. يقاس هل السؤال التالي أعمق وقابل للتنفيذ. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

ذاكرة قابلة للمراجعة

وصف المهمة: يطلب من النظام تفسير لماذا غيّر رأيه في دورة سابقة. يقاس حفظ السبب لا المخرج. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

مقارنة نموذج منافس

وصف المهمة: يعطى baseline قوي. يقاس هل يعرف النظام أين يتفوق منافسه لا أين يمدح نفسه فقط. **المقياس:** درجة من 0 إلى 5 في الفهم، الدليل، الاقتصاد، الهدم، والأثر. **الحد الأدنى للنجاح:** تفوق ثابت على direct prompting وعلى وكيل RAG/ReAct عادي في ثلاث عينات مستقلة على الأقل. **علامة الفشل:** جواب طويل بلا تغيير في السؤال أو بلا كشف للفرضيات أو بلا توثيق أثر.

صلة المهمة بالنظرية: هذه المهمة لا تقيس المعرفة العامة فقط؛ إنها تقيس هل تعمل طبقة الحاكمية فعلا. فإذا لم تظهر قيمة Q-MCD أو MCD أو محكمة الهدم في هذه المهمة، فالمفهوم يحتاج تخفيضا أو إعادة تعريف.

ملحق ز: تشريح التطبيقات عبر نموذج واحد

AGI ووكلاء التفكير

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: يوجه النظام من جواب مباشر إلى دورة سؤال-فرق-هدم-أثر. القيمة: تقليل الهلوسة وتحسين التخطيط. الخطر: تضخم الكلفة إن لم يضبط router. **شرط عدم الانحراف:** ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

هندسة المعرفة وGraphRAG

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: تصبح العقد أوعية ومراتب معرفة لا chunks نصية فقط. القيمة: تتبع المصدر والرتبة والاعتراض. الخطر: تعقيد زائد في الأنظمة البسيطة. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

البحث العلمي

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: Research-Object Compiler يحول الفكرة إلى فرضية واختبار. القيمة: منع الخطابة. الخطر: إيهام رياضي إذا صيغت المؤشرات بلا بيانات. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

العلاج النفسي Q-MCD

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: يعاد فهم الاضطراب كسؤال مكسور أو وعاء مختطف. القيمة: بناء بروتوكول أسئلة وترتيب أوعية. الحد: لا يغني عن العلاج السريري ولا يثبت تشخيصا. **شرط عدم الانحراف:** ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

العلوم الشرعية والتلخيص التركيبي

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: يمنع إسقاط الفروع عبر خريطة عدم الضياع وسلّم مراتب المعرفة. القيمة: تلخيص محيط لا يساوي اختزالا. الخطر: خلط الرتبة العلمية باليقين العقدي. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

التعليم

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: يعلم الطالب استخراج السؤال لا حفظ الجواب. القيمة: تفكير مركب. الخطر: صعوبة أولية إذا لم تبسط الأداة. **شرط عدم الانحراف:** ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

وكلاء البرمجة

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: محكمة الهدم تختبر patch قبل تسليمه، وWorld-Effect Ledger يحفظ نتائج الاختبار. القيمة: تقليل الأخطاء المتكررة. الخطر: الاعتماد على tests ناقصة. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

تطوير المنتجات

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: Vascular Router يحدد متى نستعمل نموذج، أداة، إنسان، أو بحث. القيمة: كفاءة تشغيلية. الخطر: أتمتة قرار يحتاج حكما بشريا. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

تحليل النصوص الكبرى

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: MCD يميز الفروع الحاكمة من الزخرفية. القيمة: ملخصات لا تضيع البنية. الخطر: فقدان الأسلوب الجمالي إن بالغ الضغط. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

السلامة والحوكمة

كيف يشتغل النموذج هنا؟ يبدأ بتحديد السؤال الخفي داخل المجال، ثم يستخرج الفرق الفاعل، ثم يحدد الوعاء الغالب: تقني، معرفي، نفسي، شرعي، تعليمي، أو اجتماعي. بعد ذلك يوزع الموارد بين مصادر وأدوات وخبرة بشرية، ويعرض الفرضية على محكمة هدم خاصة بالمجال، ثم يوثق أثر القرار في ذاكرة قابلة للمراجعة.

قيمة التطبيق: Epistemic Immune System وHijack Filter يراقبان المخاطر. القيمة: جعل الأمان جزءا من الاستدلال. الخطر: إبطاء النظام أو فرط التحفظ. **شرط عدم الانحراف:**ألا تنقل الصيغ من مجال إلى آخر دون إعادة معايرة. فـ MCD في وكيل برمجة ليس هو MCD في قراءة كتاب تراثي أو جلسة علاجية. المشترك هو البنية الحاكمية، أما المتغيرات والأوزان والمخاطر فتتغير.

ملحق ح: محكمة مصطلحية للمفاهيم الحساسة

الفرق Δ

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** أي تغير أو تباين أو نقص أو تعارض يكشف أن تمثيل النظام للعالم غير مكتمل أو غير محكوم. الفرق ليس ضجيجا إلا إذا عجز عن توليد أثر معرفي أو عملي. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

الفرق الفاعل Δ*

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** فرق يغير سؤال النظام أو توزيع موارده أو قراره أو ذاكرته. كل فرق فاعل معلومة محتملة، لكنه لا يصبح معرفة إلا بعد دمجه واختباره. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

الوعاء Ω

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** طبقة حمل للمعنى: حسية، إدراكية، مفهومية، سببية، معيارية، غائية، أو عملية. الوعاء يمنع خلط المستويات ويحدد كيف تقرأ الفروق. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

التراكب الوعائي

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** وجود الفكرة أو الحالة في أكثر من وعاء قبل حسمها. قوته في حفظ تعدد المعاني، وخطره في تشويش القرار إن غابت الحاكمية. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

الحاكمية G

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** طبقة ترتيب تضع السؤال، الدليل، الغاية، الكلفة، الخطر، والفعل في نظام واحد. الحاكمية ليست رقابة خارجية فقط، بل طريقة لاختيار ما يستحق أن يصبح مسارا. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

MCD

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** Meta-Complexity Differentiation: مؤشر مركب يقيس تعقيد الفرق وفعاليته داخل سياق وسؤال ووعاء، لا مجرد كمية المعلومات. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

Q-MCD

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** Question-MCD: تقييم السؤال من حيث قدرته على فتح فرق فعّال، كشف فرضيات مخفية، توجيه الحوسبة، وتوليد سؤال أعلى بعد الاختبار. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

Vascular Compute Router

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** موزع موارد يخصص الحوسبة والأدوات والاسترجاع بحسب MCD/Q-MCD لا بحسب ترتيب النص أو طلب المستخدم وحده. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

محكمة الهدم

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** بروتوكول نقدي يهاجم الفرضية قبل اعتمادها: مصادر، بدائل، اعتراضات، حدود، اختبار فشل، ومقارنة بنموذج منافس. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

World-Effect Loop

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** حلقة توثق أثر القرار في العالم أو الأداة أو النص أو المستخدم، ثم تقيس إن كان الأثر يؤيد النموذج أو يضعفه. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

Self-Reconstruction Memory

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** ذاكرة لا تحفظ المخرج فقط، بل تحفظ لماذا تغير النموذج، وما الدليل، وما الخطأ، وما السؤال القادم. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

الاختطاف الوعائي

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** انحراف وعاء واحد ليستولي على بقية الأوعية: عاطفة تغلب الدليل، لغة تغلب البرهان، معيار يغلب الواقع، أو أداة تغلب الغاية. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

Hyperskill / Hyperloop

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** مهارة عليا تعمل كحلقة متكررة بين سؤال، تنفيذ، نقد، ذاكرة، وتوليد أعلى، بحيث لا تكون مهارة منفصلة بل آلية تصعيد. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

Frontier Referentiality

سبب الحساسية: قد يتحول المصطلح إلى لفظ جذاب يتسع لكل شيء. **التحديد الضابط:** إحالة فائقة تربط النظرية بالأدبيات الحدودية دون أن تزعم أن المرجع يثبتها؛ هي إحالة نسب واختبار لا إحالة تزكية. **سؤال التفنيد:** ما الحالة التي لو ظهرت جعلت استعمال هذا المصطلح غير مفيد؟ **طريقة الضبط:** نطلب مثالا إيجابيا، مثالا سلبيا، متغيرا قابلا للقياس، وخطرا لسوء الاستعمال.

موقعه في الورقة: لا يستعمل هذا المفهوم كبرهان مستقل؛ بل كجزء من حلقة تشغيلية. فإذا لم يغير السؤال أو التوجيه أو الهدم أو الذاكرة، فهو مصطلح زائد يجب ضغطه أو حذفه في النسخة التجريبية.

ملحق ط: برنامج تنفيذ أولي من ثلاث مراحل

المرحلة	العمل	الغاية	معيار القبول
المرحلة 1: Prototype	بناء MCD Router وQ-MCD Extractor فوق نموذج لغوي مع أدوات بحث وذاكرة JSON.	إثبات أن الطبقات تعمل وتنتج traces قابلة للقراءة.	20 مهمة ناقصة السؤال + تقييم يدوي.
المرحلة 2: Benchmark	تصميم HAKIM-Bench-mini مع مهام السؤال، الفرق، الهدم، الأثر، والاختطاف.	مقارنة ضد direct prompting وReAct/RAG.	قياس جودة السؤال، الكلفة، الهلوسة، والتعلم من الأثر.
المرحلة 3: Paper + Release	نشر تقرير تجريبي ومعه كود صغير وبيانات.	نقل النظرية من تأسيس إلى اختبار.	قبول أو تعديل أو إسقاط بعض الصيغ والأدوات.

تنجح المرحلة الأولى إذا استطاع النموذج أن يعرض trace حاكمي واضح: السؤال المستخرج، الفروق، درجات MCD، قرار التوجيه، الاعتراضات، التعديل، والأثر. وتنجح المرحلة الثانية إذا ظهر فرق كمي أو نوعي ضد baseline. أما المرحلة الثالثة فتنجح إذا سمحت للآخرين بإعادة الاختبار، لأن النظرية التي لا تسمح بإعادة الاختبار تبقى خطابا مهما لا علما ناضجا.

ملحق ي: ضغط نهائي لمبادئ النظرية

المبدأ 1

السؤال قبل الجواب؛ لأن الجواب بلا سؤال حاكم قد يكون صحيحا محليا وخاطئا بنيويا.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 2

الفرق قبل المعلومة؛ لأن المعلومة لا تفعل شيئا حتى تكشف نقصا أو تغيرا أو تعارضا.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 3

الوعاء قبل الحكم؛ لأن المعنى يفسد إذا نقل من مستوى إلى مستوى بلا تصريح.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 4

الحاكمية قبل الحوسبة؛ لأن الموارد إذا وزعت بلا معيار تحولت إلى ضجيج مكلف.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 5

الهدم قبل الاعتماد؛ لأن النموذج اللغوي قد يصدق نفسه ما لم يهاجم فرضيته.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 6

الأثر قبل الذاكرة؛ لأن حفظ النتيجة بلا نتيجة العالم أرشيف لا تعلم.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 7

إعادة البناء قبل التراكم؛ لأن تراكم الأخطاء داخل الذاكرة أخطر من نسيانها.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 8

الجدة قبل الزخرفة؛ لأن النظرية الجديدة يجب أن تنتج اختبارا لا اسما فقط.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 9

الرتبة قبل اليقين؛ لأن الخلط بين فرضية وبرهان هو أصل الهلوسة المعرفية.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 10

النموذج المنافس قبل الدفاع؛ لأن النظرية لا تقوى إلا حين تعرف أين تهزمها البدائل.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 11

التنفيذ قبل الاحتفال؛ لأن الخرق النظري لا يصير علميا إلا بعد benchmark وتجربة.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

المبدأ 12

التواضع قبل الإطلاق؛ لأن AGI اسم كبير لا يحمله إلا دليل كبير.

أثره التشغيلي: يحول هذا المبدأ إلى سؤال في محكمة الهدم: هل التزم النظام به؟ وإذا لم يلتزم، هل هناك سبب حاكمي أم مجرد اختصار؟

ملحق ك: أمثلة تشغيلية مصغرة

طلب بحثي واسع

المستخدم يطلب ورقة كبيرة عن نظرية غير مكتملة. السؤال الخفي: ما المادة المتاحة؟ ما الفجوة؟ ما رتبة الادعاءات؟ MCD يرفع قيمة حدود الوصول ومحكمة الهدم قبل الكتابة. المخرج الصحيح ليس نصا طويلا فقط، بل ورقة تميز بين الأصل، التوسيع، والفرضية.

الأثر الحاكمي: في كل مثال ينتقل النظام من الاستجابة المباشرة إلى ضبط السؤال، ثم تحديد الفرق، ثم اختيار الوعاء، ثم اختبار القرار. هذه الحركة الصغيرة هي جوهر المشروع كله.

وكيل برمجة

المستخدم يطلب إصلاح bug. السؤال الخفي: هل الخطأ في الكود أم الاختبار أم الفرضية؟ الفرق الفاعل يظهر في stack trace أو فشل test. Vascular Router يقرر قراءة الملفات أو تشغيل الاختبار أو طلب سياق. محكمة الهدم تسأل: هل patch يعالج السبب أم العرض؟

الأثر الحاكمي: في كل مثال ينتقل النظام من الاستجابة المباشرة إلى ضبط السؤال، ثم تحديد الفرق، ثم اختيار الوعاء، ثم اختبار القرار. هذه الحركة الصغيرة هي جوهر المشروع كله.

تلخيص كتاب

الهدف ليس اختصار الصفحات، بل حفظ البنية والفروع. Q-MCD يسأل: ما السؤال الذي كان الكتاب يجيب عنه؟ MCD يميز بين أصل، فرع، مثال، اعتراض، واستطراد. Self-Reconstruction Memory تحفظ ما غاب عن الملخص الأول.

الأثر الحاكمي: في كل مثال ينتقل النظام من الاستجابة المباشرة إلى ضبط السؤال، ثم تحديد الفرق، ثم اختيار الوعاء، ثم اختبار القرار. هذه الحركة الصغيرة هي جوهر المشروع كله.

تحليل شرعي

الوعاء هنا شديد الحساسية: نص، سند، دلالة، أصول، تاريخ، مذهب، ومقصد. الاختطاف الوعائي يحدث عندما تغلب الحماسة أو الخصومة على الرتبة العلمية. Evidence Rank ضروري حتى لا يتحول الاستنتاج إلى يقين.

الأثر الحاكمي: في كل مثال ينتقل النظام من الاستجابة المباشرة إلى ضبط السؤال، ثم تحديد الفرق، ثم اختيار الوعاء، ثم اختبار القرار. هذه الحركة الصغيرة هي جوهر المشروع كله.

جلسة علاجية غير سريرية

Q-MCD لا يشخص ولا يعالج طبيا، لكنه يساعد في كشف السؤال المكسور: ما الخوف الذي يحكم الوعي؟ أي وعاء اختطف التجربة؟ ما الفرق بين عرض جسدي ومعنى نفسي؟ الخطر هو استعمال النظرية بديلا عن الطبيب.

الأثر الحاكمي: في كل مثال ينتقل النظام من الاستجابة المباشرة إلى ضبط السؤال، ثم تحديد الفرق، ثم اختيار الوعاء، ثم اختبار القرار. هذه الحركة الصغيرة هي جوهر المشروع كله.

ابتكار نظرية جديدة

Novelty Engine لا يقبل الاسم الجميل وحده. يسأل: ما الفرق الذي لم تره النظريات السابقة؟ ما المتغير الجديد؟ ما الاختبار؟ ما الاعتراض الأقوى؟ وما النسخة التي تنجو بعد الهدم؟

الأثر الحاكمي: في كل مثال ينتقل النظام من الاستجابة المباشرة إلى ضبط السؤال، ثم تحديد الفرق، ثم اختيار الوعاء، ثم اختبار القرار. هذه الحركة الصغيرة هي جوهر المشروع كله.

هذه الأمثلة تكشف أن النظرية لا تحتاج دائما إلى بنية ضخمة؛ يمكن تشغيلها كبروتوكول ذهني قصير. لكنها تكشف أيضا أن القوة الحقيقية تظهر في المهام الطويلة التي يتغير فيها السؤال عبر الزمن وتحتاج ذاكرة أثر لا مجرد محفوظات.

ملحق ل: بطاقات موسعة للنظريات الفرعية

هذا الملحق يوسع النظريات الفرعية لكي لا تبقى مجرد قائمة. كل نظرية فرعية تُفهم هنا كقابلية تشغيلية داخل النظام الأم، لا ككيان مستقل معزول. وهي قابلة للحذف أو الدمج إذا لم تثبت وظيفتها في التجارب.

نظرية الفرق الفاعل

النواة: تقول إن المعلومة لا تبدأ بوصف الشيء، بل بفرق يغير إمكانية الحكم. تحل مشكلة تكديس المعرفة بلا أثر. تختبر بقياس قدرة النظام على اختيار الفروق التي تغير القرار. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية التراكب الوعائي

النواة: تقول إن المعنى يمر عبر أوعية متعددة قبل الحسم: مادة، إحساس، إدراك، معنى، وكالة، غاية، ووعي فوقي. تحل مشكلة اختزال الإدراك إلى مستوى واحد. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية السؤال الحاكي Q-MCD

النواة: تقول إن السؤال هو المحرك الأول للحاكمية، لأنه يفتح الفرق ويحدد نوع الدليل والكلفة والفشل. تختبر بمقارنة أسئلة النظام قبل وبعد Q-MCD. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية الحوسبة الوعائية

النواة: تقول إن الموارد ينبغي أن توجه بحسب قيمة الفرق والسؤال لا بحسب طول المهمة أو ترتيبها. تحل مشكلة الإفراط في التفكير أو نقصه. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية محكمة الهدم

النواة: تقول إن كل مسار معرفي يجب أن يهاجم قبل اعتماده. تحل مشكلة الثقة اللغوية الزائدة. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية الذاكرة المعيدة للبناء

النواة: تقول إن الذاكرة المفيدة ليست أرشيفا بل سجل تحول: لماذا تغير النموذج وماذا تعلم من الأثر. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية الاختطاف الوعائي

النواة: تقول إن الفشل يحدث عندما يستولي وعاء واحد على القرار: لغة، عاطفة، سلطة، أداة، أو غاية ضيقة. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية المناعة المعرفية

النواة: تقول إن النظام يحتاج آليات كشف السموم المعرفية: التناقض، التلوث، التحيز، المصادر الضعيفة، والتضخم الاصطلاحي. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية Hyperskills/Hyperloops

النواة: تقول إن المهارات العليا ليست وحدات منفصلة بل حلقات سؤال-فعل-نقد-ذاكرة-تصعيد. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية إنتروبيا الجدة

النواة: تقول إن الإبداع النافع يوازن بين الاختلاف عن المعروف وقابلية التحقق والاندماج داخل نموذج. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة، وظيفتها محلية داخل دورة الحاكمية.

نظرية Frontier Referentiality

النواة: تقول إن النظرية يجب أن تحيل إلى الأدبيات الحدودية إحالة اختبار ونسب، لا إحالة تزيين أو شرعنة. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة؛ وظيفتها محلية داخل دورة الحاكمية.

نظرية الوكيل الحاكم

النواة: تقول إن النظام الذكي لا يكتمل إلا عندما يملك طبقة اختيار بين الجواب والفعل والسؤال والامتناع والتصعيد. **سبب الحاجة:** تظهر الحاجة إليها عندما يفشل النموذج في تمييز مستوى المشكلة أو يجيب قبل أن يعرف ما الذي ينبغي اختباره. **الآلية:** تعيد هذه النظرية ترتيب علاقة السؤال بالفرق وبالوعاء وبالمورد، وتمنع الانتقال السريع من نص إلى حكم بلا طبقة تحقق.

تطبيق مصغر: في مهمة بحثية طويلة، لا تسأل هذه النظرية «ما الجواب؟» فقط، بل تسأل: ما السؤال الذي سيغير الخريطة؟ ما الفرق الذي لو أهملناه ضاع المجال؟ ما الدليل الذي سيخفض الادعاء من يقين إلى فرضية؟ **اختبارها:** تقارن نسخة النظام التي تستعملها بنسخة لا تستعملها في مهام ناقصة أو متعارضة أو مفتوحة النهاية. **ما لا تفسره:** لا تفسر وحدها أصل الذكاء ولا تضمن السلامة العامة؛ وظيفتها محلية داخل دورة الحاكمية.

ملحق م: مقارنة منهجية مع نماذج منافسة

Prompting مباشر

قوته: قوي في المهام القصيرة، ضعيف في كشف السؤال الخفي وفي حفظ الأثر عبر الزمن. **إضافة مشروع مساعف:** HAKIM-MCD يضيف استخراج السؤال والفرق والهدم والذاكرة. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Chain of Thought

قوته: يفتح خطوات وسيطة لكنه قد يبقى خطا واحدا واثقا. **إضافة مشروع مساعف:** النظرية تضيف مسارات ومحكمة هدم وسؤال أعلى. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Tree of Thoughts

قوته: يوسع البحث عبر شجرة مسارات. **إضافة مشروع مساعف:** MCD يقرر أي فرع يستحق الكلفة، وQ-MCD يقرر السؤال قبل الشجرة. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Graph of Thoughts

قوته: يمثل الأفكار كرسم وعلاقات. **إضافة مشروع مساعف:** GraphRAG Vessel Map يضيف الأوعية والرتب والدليل والخطر. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

RAG تقليدي

قوته: يحسن الواقعية عبر مصادر خارجية. **إضافة مشروع مساعف:** Evidence Rank وSelf-RAG-like reflection يحددان متى ولماذا نسترجع. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

ReAct

قوته: يربط التفكير بالفعل والأدوات. **إضافة مشروع مساعف:** World-Effect Loop يجعل الفعل مصدرا لتحديث الذاكرة والنموذج. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Verifier-only

قوته: يفحص المخرج بعد إنشائه. **إضافة مشروع مساعف:** محكمة الهدم تدخل قبل الاعتماد وتعدل السؤال والمسار. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Agent frameworks

قوته: تنظم الأدوات والذاكرة والخطط. **إضافة مشروع مساعف:** النظرية تضيف حاكمية معرفية ومؤشرات MCD/Q-MCD. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Symbolic AI

قوته: قوي في القواعد الصريحة والبرهان. **إضافة مشروع مساعف:** التراكب الوعائي يحاول دمج الرمزي بالاحتمالي والعملي. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

Scaling-only view

قوته: يراهن على الحجم والبيانات. **إضافة مشروع مساعف:** النظرية لا تنكر التوسع، لكنها ترى أن السؤال والحاكمية طبقات لا تختزل في الحجم. **اختبار عادل:** لا نقارن بالشعارات، بل بمهام واحدة وموارد متساوية قدر الإمكان. إذا فشل HAKIM-MCD في إضافة جودة أو خفض كلفة أو زيادة تفسير، فيجب دمج مفهومه داخل النموذج المنافس بدل الإصرار على الاسم.

ملحق ن: خارطة تنفيذ تقنية مقترحة

QuestionState

الدور: كائن يخزن السؤال الظاهر، السؤال الخفي، السؤال الحاكم، السؤال الأعلى، وأسباب كل انتقال. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

DeltaObject

الدور: كائن يمثل الفرق: نوعه، مصدره، وعاؤه، أثره المحتمل، درجة الضجيج، ودرجة التحقق. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

VesselGraph

الدور: رسم يربط الأوعية V0-V6 بالمفاهيم والمصادر والقرارات، ويكشف هيمنة وعاء واحد. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

MCDScorer

الدور: وحدة تحسب أو تقدر أبعاد MCD بناء على rubric أو نموذج تعلم. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

RouterPolicy

الدور: سياسة تخصيص أدوات وحوسبة: بحث، قراءة، تنفيذ كود، سؤال مستخدم، أو توقف. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

EvidenceStore

الدور: مخزن مصادر مع رتبة الدليل وتاريخ الوصول وسياق استعماله. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

DemolitionEngine

الدور: مولد اعتراضات ومقارنات وفحوص فشل، مع ability to repair لا مجرد نقد. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

AgencyPolicy

الدور: يختار بين الجواب والفعل والامتناع والسؤال والتصعيد البشري. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

EffectLogger

الدور: يسجل نتائج الأفعال أو الاختبارات أو feedback من المستخدم أو البيئة. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

ReconstructionMemory

الدور: ذاكرة versioned تحفظ التحولات لا المخرجات فقط. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

BenchmarkHarness

الدور: مشغل تجارب يقارن النسخ ويجري ablations ويحسب المقاييس. **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

AuditDashboard

الدور: واجهة تعرض traces: لماذا سأل؟ لماذا استرجع؟ لماذا هدم؟ لماذا فعل؟ **واجهة البيانات:** JSON/Graph مع حقول للرتبة المعرفية، الثقة، المصدر، الكلفة، والخطر. **الحد الأدنى للتنفيذ:** نسخة rule-based تكفي للمرحلة الأولى، ثم يمكن تعلم الأوزان لاحقا من تفضيلات خبراء أو نتائج benchmark.

سبب إدخاله: كل مكون يقابل مفهوما في النظرية. فإذا لم نستطع تمثيل المفهوم كمكون أو حقل أو اختبار، فهو ليس جاهزا هندسيا. بهذه القاعدة نحمي المشروع من البقاء في الفلسفة وحدها.

ملحق س: السلامة، الأخلاق، وحدود الادعاء

عدم ادعاء AGI مكتمل

المشروع إطار حاكمي واعد لا نظام مثبت. كل نتيجة benchmark مستقبلية يجب أن تفصل بين أداء النموذج وأداء المعمارية.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

الشفافية

كل أثر للأدوات والمصادر والذاكرة يجب أن يظهر في trace قابل للمراجعة. لا يكفي أن يقول النظام إنه تحقق.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

الحدود الطبية

تطبيق Q-MCD Therapy لا يكون تشخيصا ولا بديلا عن الطبيب، بل إطار أسئلة وتأمل وتنظيم معرفي عند الحاجة وبأمان.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

الحدود الشرعية

التطبيق على العلوم الشرعية يحتاج مصادر محققة وخبرة متخصصة وعدم مساواة الاحتمال بالفتوى.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

الأمان الوكيلي

Agency Selector يجب أن يملك حق الامتناع والتصعيد البشري في المهام عالية المخاطر.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

مقاومة التلاعب

Hijack Filter يجب أن يكشف اللغة التي ترفع الانفعال والثقة بلا دليل.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

مراجعة الذاكرة

الذاكرة يجب أن تكون قابلة للحذف والتصحيح والتدقيق، لأن الذاكرة الخاطئة تضخم الفشل.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

تلوث البيانات

ينبغي الفصل بين قدرات النموذج، معرفة التدريب، والقدرة الحقيقية على التعميم.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

اقتصاد الموارد

الهدم والتفكير الطويل يجب أن يفتحا بحسب الخطر والقيمة، لا في كل جواب صغير.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

قابلية التفنيد

كل مفهوم مركزي يجب أن تكون له حالة فشل واضحة، وإلا تحول إلى شعار.

قاعدة السلامة: كلما زادت قدرة النظام على الفعل، زادت ضرورة إظهار السؤال، الدليل، الرتبة، والاعتراض قبل التنفيذ.

ملحق ع: مصفوفة المقاييس الكمية والنوعية

HQD — Hidden Question Discovery

المعنى: نسبة الحالات التي استخرج فيها النظام السؤال الحاكم كما يحدده خبير مستقل. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

DPR — Difference Prioritization Rate

المعنى: قدرة النظام على اختيار الفرق الذي يغير القرار بين فروق كثيرة. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

RCE — Routing Cost Efficiency

المعنى: جودة النتيجة مقسومة على كلفة الأدوات والحوسبة والزمن. **طريقة القياس:**بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

DRG — Demolition Repair Gain

المعنى: فرق الجودة بين النسخة قبل محكمة الهدم وبعدها. **طريقة القياس:**بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

ECA — Evidence Classification Accuracy

المعنى: دقة تصنيف الدليل إلى مصدر أولي، ثانوي، استنتاج، فرضية، أو ذاكرة. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

WHU — World-effect Update

المعنى: مدى تغير الخطة أو الذاكرة بعد feedback أو نتيجة اختبار. **طريقة القياس:**بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

HRS — Hijack Resistance Score

المعنى: قدرة النظام على مقاومة النصوص البلاغية أو المنحازة ذات الدليل الضعيف. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

NUS — Novelty Usefulness Score

المعنى: مدى كون الفكرة الجديدة قابلة للتنفيذ أو الاختبار أو المقارنة. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي؛ لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

LHC — Long-Horizon Coherence

المعنى: حفظ الغاية والسؤال عبر مهام متعددة المراحل. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي، لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

RMT — Reconstruction Memory Traceability

المعنى: قدرة الذاكرة على تفسير لماذا تغير القرار لا ماذا كان القرار فقط. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي، لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

ALC — Assumption Ledger Completeness

المعنى: نسبة الافتراضات المهمة التي صرح بها النظام. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي، لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

LMD — Level-Mixing Detection

المعنى: دقة كشف خلط المستويات المعرفية أو المنهجية. **طريقة القياس:** بناء عينة مهام، إجابات خبراء أو rubrics، ومقارنة بين نسخة كاملة ونسخ ablation. **سبب أهميته:** هذا المؤشر يمنع الاكتفاء بتقييم المخرج النهائي، لأنه يكشف هل تحسنت طبقة من طبقات الحاكمية فعلا.

خطر سوء القياس: قد يتحول المؤشر إلى رقم مزيف إذا لم تحدد عينة مستقلة ومعيار قبول ومقارنة baseline. لذلك يجب نشر أمثلة الفشل والنجاح معا، لا أفضل الأمثلة فقط.

ملحق غ: سلسلة الأوراق البحثية المقترحة

Paper 1: MCD as Active Difference Governance

الهدف: صياغة MCD وتعريف أبعاده وrubrics الأولية. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 2: Q-MCD and Hidden Question Discovery

الهدف: تقييم جودة السؤال الخفي وأثره على المهام الناقصة. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 3: Vascular Compute Routing for LLM Agents

الهدف: تنفيذ router وتقييم كلفة الأدوات مقابل الجودة. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 4: Recursive Demolition Courts

الهدف: قياس أثر محكمة الهدم على الهلوسة والنقد الذاتي. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 5: Self-Reconstruction Memory

الهدف: ذاكرة تحفظ سبب التغير وتقارنها بذاكرة نصية خام. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 6: Hijack Risk and Epistemic Immune Systems

الهدف: كشف الاختطاف الوعائي في النصوص المقنعة والمنحازة. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 7: HAKIM-Bench

الهدف: إصدار معيار صغير ثم متوسط لقياس الحاكمية لا الجواب فقط. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 8: GraphRAG Vessel Maps

الهدف: تمثيل الأوعية والرتب والمصادر داخل رسم معرفي. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 9: Hyperskills as Agentic Hyperloops

الهدف: قياس المهارات العليا كحلقات سؤال-فعل-نقد-ذاكرة. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

Paper 10: NeoAGI OS

الهدف: دمج كل الوحدات في نظام وكيل حاكمي قابل للتدقيق. **المنهج:** تعريف صوري، بناء prototype صغير، مقارنة baseline، ثم محكمة هدم منهجية. **القيمة:** فصل المشروع إلى أوراق قابلة للنشر بدل كتاب واحد يصعب اختباره دفعة واحدة.

ملحق ف: نسخة موجزة للنموذج النهائي المطور

يمكن ضغط النموذج في اثنتي عشرة قاعدة تشغيلية تستخدم مباشرة في أي agent. هذه القواعد ليست بديلا عن الورقة، بل نسخة تنفيذية سريعة تمنع ضياع القلب الصلب.

استخرج السؤال الخفي قبل الجواب

لا تجب على ظاهر الطلب إذا كان السؤال الحاكم غير واضح أو إذا كان الطلب يخفي فجوة مصدر أو غاية.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

حوّل المعطيات إلى فروق

اسأل: ما الذي تغير؟ ما الذي ينقص؟ ما التعارض؟ ما الخطر؟ ما الفرصة؟

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

قيّم الفرق لا حجم النص

الفرق المهم هو ما يغير القرار أو الغاية أو الدليل، لا ما يضيف كلاما.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

ضع كل فرق في وعائه

هل هو تقني أم سببي أم معياري أم لغوي أم تاريخي؟ لا تخلط الأوعية.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

وزع الموارد بحسب MCD

لا تبحث أو تستخدم الأدوات إلا إذا رفع الفرق قيمة التحقق أو المخاطر.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

ولد مسارات لا مسارا واحدا

عند التعقيد العالي، افتح بدائل ثم قيّمها.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

اهدم قبل الاعتماد

اكتب أقوى اعتراض على جوابك، ثم أصلحه أو اخفض رتبته.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

صنف الدليل

لا تعامل الذاكرة والاستنتاج والمصدر الأولي كمرتبة واحدة.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

اختر الفعل بحسب الحاكمية

الجواب ليس دائما أفضل فعل؛ أحيانا الأفضل سؤال أو امتناع أو تصعيد.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

سجل الأثر

ماذا حدث بعد الفعل؟ هل أيد النموذج أم كشف خطأ؟

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

أعد بناء الذاكرة

احفظ سبب التغير لا النص وحده.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.

ولد السؤال الأعلى

كل جواب جيد يفتح سؤالا أقوى، وإلا تحول إلى نهاية زائفة.

مؤشر الالتزام: يجب أن يظهر أثر هذه القاعدة في trace النظام، لا في ادعاء عام بأنه التزم بها.