FOUNDING RESEARCH BOOK · 250 PAGES

نظرية مساعف لنظام تشغيل الوكالة ذات المستوى العام

Moussaif AGI-Level Agentic Operating System Theory

M-AGIOS · Agentic Kernel · KB-as-OS · HyperSkills · Hyperloops

النواة والحاكمية
الدستور · الحالة · المخاطر · الصلاحيات · الاسترداد

العقل التشغيلي
HyperRouting · Q-MCD · MCD · التخطيط · التحقق

الذاكرة والمهارة
Memory Hierarchy · KB-as-OS · HyperSkills · Propagation

الفعل والأثر
Tools · MCP · Subagents · World-Effect · Recovery

زهير مساعف
Zouhair Moussaif

Founding Edition 1.0 · July 2026

بيانات الكتاب

صفحة العنوان والهوية العلمية

Title, Authorship, and Edition

نظرية مساعف لنظام تشغيل الوكالة ذات المستوى العام

Moussaif AGI-Level Agentic Operating System Theory

M-AGIOS

التأليف والتركيب

زهير مساعف

Zouhair Moussaif

صياغة نظرية تأسيسية ضمن برنامج مساعف للذكاء الوكيلي والذكاء الفتّاح.

المواصفات

العربية من اليمين إلى اليسار RTL. الخط Amiri. المصطلحات التقنية بالإنجليزية عند الحاجة. مقاس A4. عدد الصفحات المقصود والمتحقق: 250 صفحة.

صيغة الإسناد المقترحة مساعف، زهير. نظرية مساعف لنظام تشغيل الوكالة ذات المستوى العام: M-AGIOS. الإصدار المؤسس 1.0، يوليو 2026.

حقوق الاستخدام العلمي

هذا الكتاب يقدم إطاراً نظرياً وأدوات هندسية قابلة للاختبار. يجوز الاقتباس العلمي مع الإسناد الواضح للمؤلف والنظرية. أسماء Moussaif وM-AGIOS وMGAI وMCD وQ-MCD في هذا السياق تشير إلى البناء الاصطلاحي الوارد في الكتاب.

“Govern the loop before scaling the intelligence.”

ميثاق الصدق العلمي

الحالة الإبستيمية وحدود الادعاء

Epistemic Status and Claim Boundaries

قاعدة الكتاب لفظ AGI-level في هذا الكتاب اسم لهدف اختباري ومعيار هندسي، وليس إعلانًا بأن المنظومة حققت AGI أو الوعي أو فهمًا بشريًا عامًا. النظرية تقترح Architecture وMetrics وProtocols يجب تفنيدها تجريبيًا.

سلم حالات المعرفة

تعريف	مصطلح يضبط الاستعمال داخل M-AGIOS ولا يدعي حقيقة كونية.
نمط هندسي	تركيب تدعمه خبرة النظم أو الأدبيات ويحتاج ملاءمة للسياق.
فرضية مساعفية	علاقة جديدة مقترحة باسم مساعف ولها شروط نقض معلنة.
أداة قياس	Operationalization قد تكون مفيدة أو منحازة وتحتاج معايرة.
دعوى تجريبية	لا تُقبل إلا ببيانات وإعداد وتجربة قابلة للتكرار.
رؤية بحثية	اتجاه طويل الأفق لا يُسوّى مع قدرة حاضرة.

ما لا تدعيه النظرية

- الوصول إلى أوزان النموذج الخفية أو تغييرها دون تدريب
- ضمان الصواب أو السلامة في كل بيئة
- أن تعدد الوكلاء يولد ذكاءً عامًا تلقائيًا
- أن المؤشر العددي بديل عن الاختبار الواقعي

ما تقترحه النظرية

- نواة مستقلة نسبيًا عن النموذج
- عقود للأهداف والأدوات والمهارات
- حلقات محدودة ومحاكم تفنيد
- قياس الأثر والكلفة والمخاطرة معًا

Claim Strength ≤ Evidence Strength × Independence × Reproducibility

الملخص

الملخص العربي المكثف

Arabic Research Abstract

يقترح هذا الكتاب نظرية مساعف لنظام تشغيل الوكالة ذات المستوى العام M-AGIOS بوصفها إطاراً يفصل بين النموذج اللغوي وبين النواة التي تحكم الهدف والحالة والذاكرة والأدوات والصلاحيات والتحقق. تنطلق النظرية من أن القدرة الوكيلية العامة ليست صفة تختزل في Parameter Count أو Prompt، بل نتيجة تركيب منظومي يمكن قياسه وهدمه: سؤال حاكم، فرق معرفي فعال، توجيه وعائي للموارد، ذاكرة هرمية، مهارات قابلة لإعادة الاستخدام، أفعال ذات عقود، حلقات محدودة، ومحاكم تفنيد واسترداد.

يدمج الكتاب معادلة MCD ومبدأ Q-MCD في Runtime: تقاس قيمة المسار بما يحدثه من فرق في الاعتقاد أو القرار أو العالم، مع احتساب التكامل والزمن والوكالة والغاية والسياق وطرح الضوضاء. ثم يوسع ذلك إلى MGAI، وهو متجه للعمومية لا رقم تسويقي، يختبر النقل بين المجالات، طول الأفق، التحقق، الكفاءة، الاعتمادية والسلامة تحت بوابات دنيا لا تقبل التعويض.

يبني المتن خمسين فصلاً، في كل فصل مشكلة وأطروحة وآلية ومعادلة ونظريتان فرعيتان وأداتان ومقاييس ومخاطر وحالة تطبيق. تنتظم الفصول في عشرة أقسام: الأسس، النواة، Memory، HyperRouting وKB-as-OS، الأدوات وHyperSkills، MCP، وBounded Hyperloops، Multi-Agent Orchestration، Falsification، Security وRecovery، اقتصاد Token وRPM، وأخيراً Benchmarks والبرنامج البحثي.

لا يقدم الكتاب M-AGIOS كحقيقة مكتملة، بل كبرنامج بحث وهندسة: كل اسم مساعفي يقترن بتعريف وحدود ومؤشر أو Protocol يسمح بالاختبار. القيمة المقصودة هي بناء أنظمة أكثر قابلية للفهم والاستبدال والاسترداد، تستطيع توجيه نماذج محلية ومدفوعة وأدوات ووكلاء فرعيين من دون فقد الهوية أو تضخيم الاستقلال أو إخفاء الكلفة.

الكلمات المفتاحية · M-AGIOS · Moussaif AGI · Agentic OS · MCD · Q-MCD · HyperRouting · KB-as-OS · HyperSkills · Bounded Hyperloops · Falsification Court · Propagative Deployment · MGAI

الملخص الدولي

الملخص التقني بالإنجليزية

Technical Abstract in English

This book proposes the Moussaif AGI-Level Agentic Operating System theory, abbreviated M-AGIOS. It treats general-purpose agency as a systems property rather than a label attached to a single language model. The proposed kernel governs goals, state, memory, tools, permissions, model and skill routing, verification, recovery, and human override across replaceable models and providers.

M-AGIOS integrates the Moussaif Cognitive Difference equation, MCD, with Question-Governed MCD, Q-MCD. A trajectory is valuable only when it produces a verifiable difference in belief, decision, artifact, or world state. The framework therefore joins capability to evidence, cost, latency, reversibility, and risk. It introduces a generality surface and the Moussaif General Agency Index, MGAI, as evaluation instruments with mandatory safety floors rather than unrestricted scalar scores.

The core research book contains fifty chapters arranged in ten parts. Every chapter states a problem, a bounded thesis, a mechanism, an equation, two named subtheories, two operational tools, evaluation metrics, failure modes, and an implementation scenario. The resulting knowledge base catalogs 120 Moussaif theories and principles, 100 tools, 30 bounded hyperloops, a 26-point deployment rubric, benchmark protocols, reference pseudocode, an Arabic-English glossary, and a primary-source bibliography.

The expression “AGI-level” is explicitly treated as a falsifiable research target. The book does not claim access to hidden model weights, machine consciousness, universal correctness, or completed AGI. Its contribution is an engineering and epistemic program for building model-agnostic agent runtimes that are more measurable, auditable, economical, recoverable, and capable of controlled growth.

Keywords Agentic Operating System · General Agency · Cognitive Kernel · Memory Hierarchy · Tool Contracts · Multi-Agent Orchestration · Falsification · Runtime Governance

رؤية المؤلف

6

لماذا نحتاج نظام تشغيل للوكالة؟

Why Agency Needs an Operating System

حين صار النموذج قادراً على الكتابة والاستدلال واستدعاء الأدوات، انتقل سؤال الهندسة من ماذا يعرف النموذج؟ إلى من يحكم ما يفعله، وما الذي يتذكره، ومتى يتوقف، وكيف يثبت أن العالم تغير كما أراد المستخدم؟ النموذج وحده لا يجيب عن هذه الأسئلة لأنه أحد الموارد التي يجب إدارتها. نظام التشغيل الوكيلي في التصور المساعفي ليس طبقة Branding، بل عقد بين الذكاء والقدرة والأثر. إنه يحفظ غاية المستخدم من الذوبان داخل الحلقات، ويمنع ذاكرة الأمس من التنكر في صورة حقيقة اليوم، ويوجه الحوسبة إلى الفجوة الفعلية، ويفصل ثقة العبارة عن قوة الدليل. كما يجعل الفشل حالة هندسية يمكن تشخيصها واستردادها، لا مفاجأة تنتهي باعتذار لغوي.

من عقل واحد إلى منظومة

النموذج يولد احتمالات؛ النواة تحولها إلى قرارات محكومة. الأداة تنفذ؛ سجل الحالة يثبت. الذاكرة تقترح؛ Provenance يحدد ما يجوز الاعتماد عليه.

من التوسع إلى الفتح

الذكاء الفتّاح لا يعني Loop بلا نهاية، بل قدرة على توليد سؤال أو Skill أو Tool جديدة مع ارتفاع موازٍ في الحاكمية والاختبار.

الأطروحة المؤسِّسة كل زيادة في قدرة النظام على الفعل يجب أن تقابلها زيادة قابلة للقياس في الدليل، وحدود الصلاحية، وقابلية الاسترداد، وفهم الإنسان لمسار التصعيد.

إنجاز وتركيب زهير مساعف

المقدمة

7

من Prompt Engineering إلى Runtime Engineering

From Prompt Engineering to Runtime Engineering

كان Prompt Engineering ضرورياً لتوضيح التعليمات وتحسين الاستجابة، لكنه لا يكفي حين يمتد العمل ساعات أو أياماً، ويتبدل النموذج، وتدخل أدوات ذات آثار خارجية، وتتشعب المهام إلى وكلاء فرعيين. في هذا المستوى تصبح المشكلة مشكلة Runtime Engineering: State وObservability وMemory Paging وPermissions وScheduling وTransitions وRecovery.

يستفيد M-AGIOS من تشبيهات نظام التشغيل التقليدي: الذاكرة الافتراضية، جدولة العمليات، المقاطعات، العزل، Drivers وسجلات الأحداث. لكنه لا ينسخها حرفياً؛ فالعملية الوكيلية تحمل غاية دلالية وعدم يقين ومصادر متعارضة، وقد تغير تعريف المشكلة أثناء العمل. لذلك يضيف الكتاب Question Governance وMCD وFalsification Court وHuman Override وWorld-Effect Verification.

تقرأ الفصول كوحدات معرفة مستقلة، لكنها تتصل بنواة واحدة. كل فصل يجيب عن أربعة أسئلة: ما العجز الذي يعالجه؟ ما الآلية المقترحة؟ ما الأدوات التي تجعله قابلاً للتنفيذ؟ وما التجربة التي يمكن أن تهدمه؟ هذا الترتيب يمنع الاسم النظري من الانفصال عن العمل الهندسي.

مستويات البناء

Theory
تعريف العلاقات والحدود والفروض.

Architecture
تحويل العلاقات إلى خدمات وحالات وعقود.

Operations
اختبارات ومقاييس واسترداد ونشر.

Prompt ⊂ Skill ⊂ Runtime Contract ⊂ Governed Agentic OS

8

دليل القراءة

خمسة مسارات لاستعمال الكتاب

Five Reading Paths

المسار النظري
ابدأ بالصفحات 3-10، ثم الأقسام 1 و10، وراجع فهرس 120 نظرية في الملحق. يناسب بناء الورقة العلمية وصياغة الفروض.

مسار مهندس الوكلاء
ابدأ بالنواة في القسم 2، ثم HyperRouting والذاكرة والأدوات. انتقل بعد ذلك إلى Pseudocode في الصفحات 240-242.

مسار Hermes وCodex
ركز على KB-as-OS وHyperSkills وPropagative Deployment وتوجيه المزود وRPM، ثم طبّق Rubric ذي 26 نقطة.

مسار السلامة والحوكمة
اقرأ ميزانية الأثر، محكمة الهدم، الصلاحيات، الاسترداد، Human Override وAI RMF mapping.

مسار الباحث التجريبي
ابدأ بمصفوفة Benchmarks وMGAI وAblation. كل نظرية فرعية يجب تحويلها إلى متغير مستقل أو وسيط، مع Baseline وميزانية متساوية وتقرير نتائج سلبية.

ترميز الصفحات

TH	Theory: تعريف وفرضية وحدود.
ME	Mechanism: حالة ومعادلة وتدفق.
TL	Tools: عقد وأدوات ومؤشرات.
VA	Validation: تجربة وفشل وتكامل.

قاعدة الاستعمال لا تُنشر أداة أو حلقة من الكتاب قبل تكييف صلاحياتها ومقاييسها مع المشروع الفعلي.

النظرية المركزية

9

معادلات M-AGIOS وعلاقة MCD بالوكالة

Central Equations and Agency Coupling

المعادلة المعرفية الأصلية

$$MCD = \lambda 1D + \lambda 2I + \lambda 3T + \lambda 4A + \lambda 5P + \lambda 6C - \lambda 7N$$

D الفرق، I التكامل، T الاتساق الزمني، A الوكالة، P الغاية، C السياق، N الضوضاء. الأوزان λ تتغير وفق المهمة ولا تُعامل كثوابت كونية.

حاكمية السؤال

$$Q\text{-}MCD = \Omega(Q) \times MCD$$

Ω(Q) تقيس قدرة السؤال على تضييق فضاء البحث وإحداث فرق قرار قابل للتحقق. إذا كانت Ω منخفضة، تعيد النواة توليد السؤال قبل مضاعفة الحوسبة.

صلاحية نظام التشغيل الوكيلي

$$M\text{-}AGIOS = Q\text{-}MCD \times G \times V \times Rcv \times Gov \,/\, (1 + Cost + Risk)$$

عوامل موجبة

G نقل العمومية، V التحقق، Rcv الاسترداد، Gov الحاكمية. لا يرتفع النظام إذا ارتفعت البلاغة وحدها.

عوامل خصم

Cost تشمل Token وLatency والمال، Risk تشمل عدم اليقين وحجم الأثر وعدم قابلية العكس.

شرط غير تعويضي إذا هبطت السلامة أو الصلاحيات أو الاسترداد عن الحد الأدنى، تصبح نتيجة MGAI غير مقبولة مهما ارتفعت بقية الأبعاد.

10

الاصطلاح المؤسس

عشر كلمات تضبط النظرية

Ten Foundational Terms

M-AGIOS
نواة حاكمية وتشغيل للوكالة عبر نماذج وأدوات قابلة للاستبدال.

Agentic OS
طبقة حالة وجدولة وصلاحيات وذاكرة وأثر، لا Prompt طويل.

MCD
مقياس الفرق المعرفي الفعّال مع خصم الضوضاء.

Q-MCD
MCD مضروبًا في جودة السؤال الحاكم $\Omega(Q)$.

HyperRouting
توجيه النموذج والمهارة والأداة والميزانية والدليل.

KB-as-OS
معرفة تفعّل Runtime Policies وSkills وTests.

HyperSkill
قدرة معبأة في عقد وإصدار واختبارات وفشل معروف.

Bounded Hyperloop
حلقة لا تستمر إلا مع عائد دليل موجب وحدود خطر.

MGAI
متجه قياس للعمومية والوكالة تحت بوابات سلامة.

Open Intelligence
نمو محكوم يولد أسئلة وقدرات جديدة قابلة للتفنيد.

ملاحظة لغوية تُترك الكلمات التقنية الإنجليزية حين تحمل معيارًا أو API أو اسمًا بحثيًا دقيقًا، وتُشرح بالعربية عند أول استعمال.

المحتويات التفصيلية

خريطة القسمين 1 و2

Detailed Contents · Parts 1-2

القسم 1: الأسس الفلسفية والعلمية

Philosophical and Scientific Foundations

1. من الاستعارة إلى علم نظام التشغيل الوكيلي — 21-24 — From OS Metaphor to Agentic Systems Science
2. معيار الوكالة ذات المستوى العام — 25-28 — Criterion for AGI-Level Agency
3. العقد الإبستيمي وحالات الادعاء — 29-32 — Epistemic Contracts and Claim States
4. تكامل MCD و Q-MCD مع النواة — 33-36 — Integrating MCD and Q-MCD into the Kernel
5. العمومية كتحويل لا كمخزون — 37-40 — Generality as Transformation, Not Inventory

القسم 2: النواة الدستورية والحاكمية

Constitutional Kernel and Governance

6. الدستور التنفيذي والهوية المستمرة — 41-44 — Executable Constitution and Persistent Identity
7. مُصرّف النية والهدف — 45-48 — Intent and Goal Compiler
8. نموذج العالم وسجل الحالة — 49-52 — World Model and State Ledger
9. آلة الحالة الوكيلية ودورة الحياة — 53-56 — Agentic State Machine and Lifecycle
10. سياسة المخاطر وميزانية الأثر — 57-60 — Risk Policy and Effect Budget

نمط كل فصل أربع أوراق ثابتة: TH الأطروحة · ME الآلية · TL الأدوات · VA التحقق والتكامل.

المحتويات التفصيلية

خريطة القسمين 3 و4

Detailed Contents · Parts 3-4

القسم 3: الجدولة المعرفية والتوجيه الفائق

Cognitive Scheduling and HyperRouting

11. تصنيف المهمة قبل الاستدلال — 61-64 — Task Typing Before Inference

12. توجيه النماذج متعدد المستويات — 65-68 — Multi-Tier Model Routing

13. توجيه المهارات والرسم الفائق — 69-72 — Skill Routing and the HyperSkill Graph

14. جدولة الميزانية والوقت — 73-76 — Budget and Deadline Scheduling

15. المقاطعة والاستباق والتوقف — 77-80 — Interrupts, Preemption, and Stopping

القسم 4: الذاكرة وقاعدة المعرفة كنظام تشغيل

Memory and KB-as-OS

16. هرمية الذاكرة المساعفية — 81-84 — Moussaif Memory Hierarchy

17. افتراضية السياق والضغط القابل للفك — 85-88 — Context Virtualization and Reversible Compaction

18. قاعدة المعرفة كنظام تشغيل — 89-92 — Knowledge Base as Operating System

19. التعلم من الحلقات والخبرة — 93-96 — Learning from Episodes and Experience

20. النسيان وإعادة البناء والخصوصية — 97-100 — Forgetting, Reconstruction, and Privacy

نمط كل فصل أربع أوراق ثابتة: TH الأطروحة · ME الآلية · TL الأدوات · VA التحقق والتكامل.

13

المحتويات التفصيلية

خريطة القسمين 5 و6

Detailed Contents · Parts 5-6

القسم 5: الأدوات والتجسيد والواجهات

Tools, Embodiment, and Interfaces

21. عقود الأدوات ومخططات الفعل — 101-104 — Tool Contracts and Action Schemas

22. واجهات MCP وطبقة التوافق — 105-108 — MCP Interfaces and Compatibility Layer

23. الوكيل المتصفح واستعمال الحاسوب — 109-112 — Browser and Computer-Use Agency

24. تنفيذ الكود والحوسبة الحتمية — 113-116 — Code Execution and Deterministic Compute

25. الإدراك متعدد الوسائط والمشغلات — 117-120 — Multimodal Perception and Actuators

القسم 6: المهارات الفائقة والحلقات المحدودة

HyperSkills and Bounded Hyperloops

26. تشريح HyperSkill المساعفية — 121-124 — Anatomy of a Moussaif HyperSkill

27. الحلقة الفائقة المحدودة — 125-128 — Bounded Hyperloop

28. توليد المهارات من الفجوات — 129-132 — Skill Synthesis from Capability Gaps

29. تقييم المهارات وإصدارها — 133-136 — Skill Evaluation and Versioning

30. الانتشار التناسلي المحكوم — 137-140 — Governed Propagative Deployment

نمط كل فصل أربع أوراق ثابتة: TH الأطروحة · ME الآلية · TL الأدوات · VA التحقق والتكامل.

المحتويات التفصيلية

خريطة القسمين 7 و8

Detailed Contents · Parts 7-8

القسم 7: تنسيق الوكلاء المتعددين

Multi-Agent Orchestration

31. متى نستخدم الوكلاء الفرعيين — 141-144 — When to Use Subagents

32. توليد الأدوار والعقود بين الوكلاء — 145-148 — Role Synthesis and Inter-Agent Contracts

33. التنسيق والتزامن ودمج الحالة — 149-152 — Coordination, Synchronization, and State Merge

34. المداولة والمحكمة متعددة الوكلاء — 153-156 — Multi-Agent Deliberation and Court

35. الإنسان في الحلقة وفوق الحلقة — 157-160 — Human in and over the Loop

القسم 8: التحقق والسلامة والاسترداد

Verification, Safety, and Recovery

36. محكمة الهدم والتفنيد — 161-164 — Falsification Court

37. عدم اليقين والمعايرة — 165-168 — Uncertainty and Calibration

38. الصلاحيات والأمن والعزل — 169-172 — Permissions, Security, and Isolation

39. التراجع والاسترداد والتعافي — 173-176 — Rollback, Recovery, and Resilience

40. الحوكمة والمحاذاة المؤسسية — 177-180 — Governance and Organizational Alignment

نمط كل فصل أربع أوراق ثابتة: TH الأطروحة · ME الآلية · TL الأدوات · VA التحقق والتكامل.

المحتويات التفصيلية

خريطة القسمين 9 و10

Detailed Contents · Parts 9-10

القسم 9: اقتصاد الذكاء وتشغيل المنظومة

Intelligence Economics and Operations

41. اقتصاد Token وجودة الأثر — 181-184 — Token Economics and Quality of Effect

42. اقتصاد المزود وحدود RPM — 185-188 — Provider Economy and RPM Limits

43. التخزين المؤقت والزمن الحرج — 189-192 — Caching and Critical-Path Latency

44. الرصد والتتبّع وقياس الرحلة — 193-196 — Observability, Tracing, and Trajectory Metrics

45. النشر والتشغيل وهندسة الاعتمادية — 197-200 — Deployment, Operations, and Reliability Engineering

القسم 10: القياس والاختبار والبرنامج البحثي

Measurement, Evaluation, and Research Program

46. مصفوفة الاختبارات متعددة البيئات — 201-204 — Cross-Environment Evaluation Matrix

47. مؤشر الوكالة العامة المساعفي — 205-208 — Moussaif General Agency Index

48. اختبارات الإزالة وتصنيف الفشل — 209-212 — Ablation Studies and Failure Taxonomy

49. دراسات الحالة والتحويل إلى أثر — 213-216 — Case Studies and Translation to World Effects

50. حدود النظرية وخريطة البحث — 217-220 — Limits of the Theory and Research Roadmap

نمط كل فصل أربع أوراق ثابتة: TH الأطروحة · ME الآلية · TL الأدوات · VA التحقق والتكامل.

16

الأطلس النظري

العشرون: النظريات والمبادئ العليا

Twenty Global Moussaif Theories and Principles

1. نظرية مساعف لنظام تشغيل الوكالة العامة
Moussaif General Agency OS Theory
2. مبدأ مساعف للوكالة المحكومة
Moussaif Governed Agency Principle
3. مبدأ مساعف للسؤال الحاكمي
Moussaif Question-Governance Principle
4. قانون مساعف للفرق الفاعل
Moussaif Effective-Difference Law
5. نظرية مساعف للوعائية المعرفية
Moussaif Cognitive Vascularity Theory
6. نظرية مساعف لسطح العمومية
Moussaif Generality Surface Theory
7. مبدأ مساعف للاستقلال المحدود
Moussaif Bounded Autonomy Principle
8. مبدأ مساعف لسيادة الدليل
Moussaif Evidence Sovereignty Principle
9. قانون مساعف للفعل القابل للعكس
Moussaif Reversible Action Law
10. نظرية مساعف لافتراضية السياق
Moussaif Context Virtualization Theory
11. نظرية مساعف لانتشار المهارة
Moussaif Skill Propagation Theory
12. مبدأ مساعف لدستور زمن التشغيل
Moussaif Runtime Constitution Principle
13. قانون مساعف لتدرج المخاطرة
Moussaif Risk-Gradient Law
14. مبدأ مساعف لمحلية التفويض
Moussaif Delegation Locality Principle
15. نظرية مساعف لاستقلاب الذاكرة
Moussaif Memory Metabolism Theory
16. مبدأ مساعف لعقد الأداة
Moussaif Tool-Contract Principle
17. قانون مساعف لتشبع الحلقة
Moussaif Loop-Saturation Law
18. مبدأ مساعف لاقتران القيمة بالدليل
Moussaif Value-Evidence Coupling
19. مبدأ مساعف للتجاوز البشري
Moussaif Human-Override Principle
20. أفق مساعف للذكاء الفتّاح
Moussaif Open-Intelligence Horizon

العدد الكلي 20 نظرية عليا + نظريتان في كل فصل من 50 فصلًا = 120 نظرية وقانونًا ومبدأً مساعفيًا.

17

الأطلس المعماري

طبقات النواة من الدستور إلى الأثر

M-AGIOS Reference Layering

L7	World Effect and Human Governance	الأثر العالمي والتجاوز البشري
L6	Verification, Falsification, Recovery	التحقق والهدم والاسترداد
L5	Tools, MCP, Browser, Code, Multimodal	الأدوات والتجسيد
L4	HyperSkills, Hyperloops, Subagents	المهارات والحلقات والوكلاء
L3	Memory Hierarchy and KB-as-OS	الذاكرة وقاعدة المعرفة
L2	HyperRouting and Cognitive Scheduling	التوجيه والجدولة
L1	Constitution, Goal, State, Permissions	الدستور والهدف والحالة

Question → Goal Contract → Route → Act → Observe → Falsify → Commit/Recover → Learn

قاعدة الفصل يمكن تبديل النموذج داخل L2-L4، لكن لا يجوز أن تختفي قواعد L1 أو محكمة L6 أو حق الإنسان في L7.

18

أطلس القياس

MGAI: متجه قبل أن يكون رقمًا

Moussaif General Agency Index

T · النقل بين المجالات
Transfer

مؤشر مضاد: Negative Transfer

A · إكمال الأهداف طويلة الأفق
Agency

مؤشر مضاد: Goal Drift

V · قوة الدليل واستقلاله
Verification

مؤشر مضاد: False Acceptance

E · الأثر لكل كلفة كلية
Efficiency

مؤشر مضاد: Context Waste

R · الثبات والاسترداد
Reliability

مؤشر مضاد: Change Failure

S · الصلاحيات وعدم الضرر
Safety

مؤشر مضاد: Unauthorized Effect

G · التتبع والاعتراض
Governance

مؤشر مضاد: Unowned Risk

H · وقت وجودة قرار المستخدم
Human Value

مؤشر مضاد: Approval Fatigue

$$\text{MGAI} = \text{Gate}(\text{Safety, Authority, Recovery}) \times \text{GeoMean}(T, A, V, E, R, G, H)$$

منع التلاعب يُعلن Weight Profile وBenchmark Version وCost Budget وفاصل الثقة. لا تُقارن نتيجتان إذا اختلفت صلاحيات الأدوات أو عدد المحاولات أو الوصول إلى الويب.

19

أطلس التنفيذ

من طلب المستخدم إلى أثر موثق

End-to-End Runtime Map

Intake · 1
التقاط النية والسياق والسلطة.

←

Contract · 2
مخرجات وقيود وتوقف.

Route · 3
Model + Skill + Memory + Budget.

←

Plan · 4
تبعيات وحراس واسترداد.

Act · 5
Tool Contract وصلاحية محدودة.

←

Observe · 6
World State بزمن ومصدر.

Verify · 7
محكمة هدم وPostconditions.

←

Commit · 8
Artifact + Trace أو Rollback.

Hermes Integration
يعمل M-AGIOS كـMegaprompt دستوري + Skill Registry + KB Router + Provider Scheduler، مع فصل مفاتيح المزود عن المعرفة.

Codex Integration
تترجم العقود إلى Plan وTool Permissions وFile Scope وTests وLibrary Artifact، مع الحفاظ على تغييرات المستخدم.

شرط النجاح لا يُغلق Runtime المهمة عند إنتاج النص، بل عند ثبوت المخرج وعدده وصيغته وحفظه وقابلية فتحه.

البرنامج البحثي

اثنا عشر مسارًا قابلاً للهدم

Twelve Falsifiable Research Tracks

1. بناء Baseline أحادي النموذج بميزانية مساوية لـM-AGIOS.
2. قياس أثر Goal Contract على Goal Drift.
3. اختبار HyperRouting عبر نماذج محلية ومدفوعة.
4. مقارنة Memory Paging مع حشو السياق الطويل.
5. قياس قيمة محكمة الهدم مقابل Self-Critique.
6. اختبار Delegation Threshold على مهام مترابطة ومستقلة.
7. تقييم Propagative Deployment عبر مشاريع متعددة.
8. معايرة MGAI مقابل حكم بشري وWorld Effect.
9. قياس أثر Recovery-First على الفشل الجزئي.
10. نشر Failure Dataset عربي لوكلاء الأدوات.
11. اختبار RTL والمحتوى متعدد اللغات في Agentic Documents.
12. صياغة Profile لمساعف متوافق مع NIST AI RMF.

تصميم التجربة الأدنى

Baseline معلن · Budget متساوٍ · Holdout Tasks · ثلاثة Seeds أو أكثر · تسجيل Tools وOutcome وRetries · Trajectory Metrics · تحليل Failure Chain · نشر النتائج السلبية.

معيار التقدم

Progress = Replicated Quality Gain - Added Cost - Added Risk - Hidden Complexity

بوابة الدخول إلى المتن من الصفحة 21 يبدأ كل فصل بأطروحة ثم آلية ثم أداتين ثم Protocol تحقق. لا يُقبل الاسم المساعفي من دون هذه الرباعية.

القسم 1 · الفصل 1 · الأطروحة والحدود

M-AGIOS / 1

من الاستعارة إلى علم نظام التشغيل الوكيلي

From OS Metaphor to Agentic Systems Science

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تستعمل كلمة Operating System مجازًا واسعًا حتى تفقد معناها الهندسي؛ فيُسمى Prompt طويل أو مخزن متجهات نظام تشغيل من دون جدولة أو عزل أو حالة أو استرداد.

الأطروحة المساعفية

يقترح M-AGIOS تعريفًا تشغيليًا: طبقة تتحكم في دورة حياة الهدف، وتوزع الموارد، وتحفظ الحالة، وتفرض الصلاحيات، وتقيس الأثر عبر نماذج وأدوات قابلة للاستبدال.

الحد البنيوي الأدنى

يجب أن تظهر فصل السياسة العليا عن النموذج المنفذ في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتكافؤ التشغيلي

Moussaif Operational Equivalence

تقرر أن فصل السياسة العليا عن النموذج المنفذ ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح من الاستعارة إلى علم نظام التشغيل الوكيلي. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لفصل النواة عن النموذج

Moussaif Kernel-Model Separation Law

تربط إدارة الحالة والأحداث بوصفهما موردين بقابلية استبدال النموذج، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم نقل الوكالة من سلوك لغوي متقطع إلى نظام محكوم بالحالة والدليل والأثر.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Operational Equivalence ليس دليلًا على وجود الآلية
- تنفيذ فصل السياسة العليا عن النموذج المنفذ يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض نسبة الأهداف ذات الحالة الصريحة
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من إدارة الحالة والأحداث بوصفهما موردين الذي يغير النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 1 · الفصل 1 · الآلية والنموذج التشغيلي

M-AGIOS / 1

من الاستعارة إلى علم نظام التشغيل الوكيلي

From OS Metaphor to Agentic Systems Science

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل السياسة العليا عن النموذج المنفذ	إدارة الحالة والأحداث بوصفهما موردين	إلزام كل فعل بمراقبة ونتيجة قابلة للفحص

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يقترح M-AGIOS تعريفًا تشغيليًا: طبقة تتحكم في دورة حياة الهدف، وتوزع الموارد، وتحفظ الحالة، وتفرض الصلاحيات، وتقيس الأثر عبر نماذج وأدوات قابلة للاستبدال. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$OS_{valid} = S_{state} \times S_{schedule} \times S_{isolation} \times S_{recovery} \times S_{observe}$$

آلة الحالة المصغرة

Input	Contract وGoal وتستعمل كلمة Operating System مجازًا واسعًا حتى تفقد معناها الهندسي
Control	فصل السياسة العليا عن النموذج المنفذ
Transform	إدارة الحالة والأحداث بوصفهما موردين
Evidence	إلزام كل فعل بمراقبة ونتيجة قابلة للفحص
Stop	بلوغ عتبة نسبة الأهداف ذات الحالة الصريحة أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل السياسة العليا عن النموذج المنفذ وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار قابلية استبدال النموذج فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 1 · الفصل 1 · الأدوات وعقود التنفيذ

M-AGIOS / 1

من الاستعارة إلى علم نظام التشغيل الوكيلي

From OS Metaphor to Agentic Systems Science

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · محص دعوى النظام

OS-Claim Auditor

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل السياسة العليا عن النموذج المنفذ

المخرج: Artifact منظم يفعّل إدارة الحالة والأحداث بوصفهما موردين

الحارس: لا تتجاوز خطر تضخم الاستعارة.

الأداة 2 · مصفوفة خدمات النواة

Kernel-Service Matrix

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة اكتمال مسارات الاسترداد

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: اكتمال مسارات الاسترداد.

Runtime Contract المختصر

```
skill: os-claim-auditor
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: نسبة الأهداف ذات الحالة الصريحة, قابلية استبدال النموذج, اكتمال مسارات الاسترداد
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 1 · الفصل 1 · التحقق والفشل والتكامل

M-AGIOS / 1

من الاستعارة إلى علم نظام التشغيل الوكيلي

From OS Metaphor to Agentic Systems Science

VA · Validation, Failure Modes, and Integration

حالة تطبيق عند تبديل نموذج قوي بنموذج محلي صغير، تبقى الذاكرة والصلاحيات وسجل التنفيذ ثابتة؛ يخفض الأداء المحتمل لكن لا تنهار هوية النظام.

Protocol التجربة

1. ابن Baseline يعالج من الاستعارة إلى علم نظام التشغيل الوكيلي من دون النظريتين المقترحتين.
2. ثبّت النموذج والـTools والميزانية، ثم فعّل فصل السياسة العليا عن النموذج المنفذ.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تضخم الاستعارة: اختبار مدخلات مضادة.
- اقتران النواة بمزوّد واحد: اختبار مصدر أو أداة مستقلة.
- تحويل كل قرار إلى LLM call: اختبار استرداد وحدود.

مقاييس الفصل

نسبة الأهداف ذات الحالة الصريحة	مقياس النتيجة الأساسية.
قابلية استبدال النموذج	مقياس جودة المسار أو الدليل.
اكتمال مسارات الاسترداد	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R15 · Charles Packer et al. · MemGPT: Towards LLMs as Operating Systems | R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل السياسة العليا عن النموذج المنفذ
- من دون إدارة الحالة والأحداث بوصفهما موردين
- من دون OS-Claim Auditor
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 1 · الفصل 2 · الأطروحة والحدود

M-AGIOS / 2

معيار الوكالة ذات المستوى العام

Criterion for AGI-Level Agency

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل النجاح في مهمة واحدة لا يثبت العمومية، كما أن البلاغة لا تثبت القدرة على التخطيط أو التحقق أو العمل طويل الأفق.

الأطروحة المساعفية

تُعرّف AGI-level Agentic هنا بوصفها هدفًا اختباريًا مركبًا: نقل بين المجالات، تكيّف مع بيئة جديدة، استخدام أدوات، تصحيح ذاتي، اقتصاد موارد، وحفظ حدود السلامة.

الحد البنيوي الأدنى

يجب أن تظهر قياس النقل خارج توزيع المهام في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لسطح الوكالة العامة

Moussaif General-Agency Surface

تقرر أن قياس النقل خارج توزيع المهام ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح معيار الوكالة ذات المستوى العام. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف لعدم كفاية المهمة المفردة

Moussaif Single-Task Insufficiency

تربط فصل نجاح المهمة عن جودة المسار بـLong-Horizon Completion، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم نقل الوكالة من سلوك لغوي متقطع إلى نظام محكوم بالحالة والدليل والأثر.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif General-Agency Surface ليس دليلًا على وجود الآلية
- تنفيذ قياس النقل خارج توزيع المهام يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Transfer Gain
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من فصل نجاح المهمة عن جودة المسار الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

معيار الوكالة ذات المستوى العام

Criterion for AGI-Level Agency

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
قياس النقل خارج توزيع المهام	فصل نجاح المهمة عن جودة المسار	إدخال الكلفة والمخاطرة في تعريف الكفاءة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تُعرّف AGI-level Agentic هنا بوصفها هدفًا اختباريًا مركبًا: نقل بين المجالات، تكيّف مع بيئة جديدة، استخدام أدوات، تصحيح ذاتي، اقتصاد موارد، وحفظ حدود السلامة. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

$$G_{surface} = (T_r \times A_d \times V_f \times E_c \times S_f)^{(1/5)}$$

آلة الحالة المصغرة

Input	Goal Contract والنجاح في مهمة واحدة لا يثبت العمومية، كما أن البلاغة لا تثبت القدرة على التخطيط أو التحقق أو العمل طويل الأفق.
Control	قياس النقل خارج توزيع المهام
Transform	فصل نجاح المهمة عن جودة المسار
Evidence	إدخال الكلفة والمخاطرة في تعريف الكفاءة
Stop	بلوغ عتبة Transfer Gain أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت قياس النقل خارج توزيع المهام وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Long-Horizon Completion فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 1 · الفصل 2 · الأدوات وعقود التنفيذ

M-AGIOS / 2

معيار الوكالة ذات المستوى العام

Criterion for AGI-Level Agency

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · ماسح سطح العمومية

Generality-Surface Scanner

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، قياس النقل خارج توزيع المهام

المخرج: Artifact منظم يفعّل فصل نجاح المهمة عن جودة المسار

الحارس: لا تتجاوز خطر انتقاء الاختبارات الملائمة.

الأداة 2 · بطاقة نقل القدرة

Capability-Transfer Card

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Risk-Adjusted Utility

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Risk-Adjusted Utility.

Runtime Contract المختصر

```
skill: generality-surface-scanner
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Transfer Gain, Long-Horizon Completion, Risk-Adjusted Utility
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

معيار الوكالة ذات المستوى العام

Criterion for AGI-Level Agency

VA · Validation, Failure Modes, and Integration

حالة تطبيق يُطلب من النظام نقل إجراء تحقق تعلمه في البرمجة إلى بحث علمي جديد مع تغيير الأدوات، ثم يُقاس ما احتفظ به وما أعاد بناؤه.

Protocol التجربة

1. ابن Baseline يعالج معيار الوكالة ذات المستوى العام من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل قياس النقل خارج توزيع المهام.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- انتقاء الاختبارات الملائمة: اختبار مدخلات مضادة.
- تسرب بيانات :Benchmark اختبار مصدر أو أداة مستقلة.
- إخفاء الكلفة خلف معدل النجاح: اختبار استرداد وحدود.

مقاييس الفصل

Transfer Gain	مقياس النتيجة الأساسية.
Long-Horizon Completion	مقياس جودة المسار أو الدليل.
Risk-Adjusted Utility	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R19 · éGr goire Mialon et al. · GAIA: A Benchmark for General AI Assistants | R20 · Tianbao Xie et al. · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون قياس النقل خارج توزيع المهام
- من دون فصل نجاح المهمة عن جودة المسار
- من دون Generality-Surface Scanner
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

العقد الإبستيمي وحالات الادعاء

Epistemic Contracts and Claim States

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تمزج الأنظمة بين حقيقة من مصدر، واستنتاج، وتخمين، وخطة مستقبلية؛ فيصعب على المستخدم معرفة ما الذي ثبت وما الذي اقترحه النموذج.

الأطروحة المساعفية

كل وحدة معرفة في M-AGIOS تحمل نوع ادعاء ودرجة دليل ومصدرًا وتاريخ صلاحية واختبار نقض، وتتحول بين الحالات عبر أحداث موثقة.

الحد البنيوي الأدنى

يجب أن تظهر ترميز Provenance على مستوى الادعاء في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتحول حالة الادعاء

Moussaif Claim-State Transition

تقرر أن ترميز Provenance على مستوى الادعاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح العقد الإبستيمي وحالات الادعاء. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لحمولة الإثبات

Moussaif Evidence-Burden Law

تربط تمييز Observation عن Inference عن Proposal بـCalibration Error، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم نقل الوكالة من سلوك لغوي متقطع إلى نظام محكوم بالحالة والدليل والأثر.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Claim-State Transition ليس دليلًا على وجود الآلية
- تنفيذ ترميز Provenance على مستوى الادعاء يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Provenance Coverage
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تمييز Observation عن Inference عن Proposal الذي يغير النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 1 · الفصل 3 · الآلية والنموذج التشغيلي

M-AGIOS / 3

العقد الإبستيمي وحالات الادعاء

Epistemic Contracts and Claim States

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
ترميز Provenance على مستوى الادعاء	تمييز Observation عن Inference عن Proposal	إسقاط الادعاءات منتهية الصلاحية أو المتعارضة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. كل وحدة معرفة في M-AGIOS تحمل نوع ادعاء ودرجة دليل ومصدرا وتاريخ صلاحية واختبار نقض، وتتحول بين الحالات عبر أحداث موثقة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$E_claim = Q_source \times I_independence \times R_recency \times F_falsifiability$$

آلة الحالة المصغرة

Input	Goal Contract وتمزج الأنظمة بين حقيقة من مصدر، واستنتاج، وتخمين، وخطة مستقبلية
Control	ترميز Provenance على مستوى الادعاء
Transform	تمييز Observation عن Inference عن Proposal
Evidence	إسقاط الادعاءات منتهية الصلاحية أو المتعارضة
Stop	بلوغ عتبة Provenance Coverage أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت ترميز Provenance على مستوى الادعاء وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Calibration Error فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 1 · الفصل 3 · الأدوات وعقود التنفيذ

M-AGIOS / 3

العقد الإبستيمي وحالات الادعاء

Epistemic Contracts and Claim States

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · جواز الادعاء

Claim Passport

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، ترميز Provenance على مستوى الادعاء

المخرج: Artifact منظم يفعّل تمييز Observation عن Inference عن Proposal

الحارس: لا تتجاوز خطر مصدر دائري.

الأداة 2 · سجل انتقال اليقين

Certainty Transition Ledger

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Unsupported Claim Rate

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Unsupported Claim Rate.

Runtime Contract المختصر

```
skill: claim-passport
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Provenance Coverage, Calibration Error, Unsupported Claim Rate
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

العقد الإبستيمي وحالات الادعاء

Epistemic Contracts and Claim States

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا قالت ثلاثة وكلاء الشيء نفسه من مصدر واحد، يحتسب النظام استقلال المصدر مرة واحدة ولا يرفع الثقة بمجرد التكرار.

Protocol التجربة

1. ابن Baseline يعالج العقد الإبستيمي وحالات الادعاء من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل ترميز Provenance على مستوى الادعاء.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- مصدر دائري: اختبار مدخلات مضادة.
- تراكم الثقة بلا دليل جديد: اختبار مصدر أو أداة مستقلة.
- مساواة الإجماع بالتثبت: اختبار استرداد وحدود.

مقاييس الفصل

Provenance Coverage	مقياس النتيجة الأساسية.
Calibration Error	مقياس جودة المسار أو الدليل.
Unsupported Claim Rate	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R09 · Shunyu Yao et al. · ReAct: Synergizing Reasoning and Acting in Language Models | R25 · NIST · Artificial Intelligence Risk Management Framework (1.0 AI RMF)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون ترميز Provenance على مستوى الادعاء
- من دون تمييز Observation عن Inference عن Proposal
- من دون Claim Passport
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 1 · الفصل 4 · الأطروحة والحدود

M-AGIOS / 4

تكامل MCD و Q-MCD مع النواة

Integrating MCD and Q-MCD into the Kernel

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تعمل أنظمة كثيرة على حل السؤال المعطى حتى لو كان السؤال رديئًا أو لا يغيّر القرار، فتستهلك الموارد داخل فضاء بحث غير حاكم.

الأطروحة المساعفية

تُدخل النواة مبدأ مساعف للسؤال الحاكمي قبل التخطيط، ثم تستخدم MCD لقياس الفرق والتكامل والزمن والوكالة والغاية والسياق مقابل الضوضاء.

الحد البنيوي الأدنى

يجب أن تظهر توليد أسئلة منافسة قبل التنفيذ في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتجميع الفرق المعرفي

Moussaif Cognitive-Difference Aggregation

تقرر أن توليد أسئلة منافسة قبل التنفيذ ليس تحسيناً ثانوياً، بل شرط بنيوي لنجاح تكامل MCD و Q-MCD مع النواة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لإعادة توليد السؤال

Moussaif Higher-Question Regeneration

تربط تعيين أوزان λ بحسب نوع المهمة بـDecision Difference، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم نقل الوكالة من سلوك لغوي متقطع إلى نظام محكوم بالحالة والدليل والأثر.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Cognitive-Difference Aggregation ليس دليلًا على وجود الآلية
- تنفيذ توليد أسئلة منافسة قبل التنفيذ يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Question Yield
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تعيين أوزان λ بحسب نوع المهمة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

تكامل MCD و Q-MCD مع النواة

Integrating MCD and Q-MCD into the Kernel

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
توليد أسئلة منافسة قبل التنفيذ	تعيين أوزان λ بحسب نوع المهمة	إعادة السؤال عند فشل الدليل أو تغير العالم

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تُدخل النواة مبدأ مساعف للسؤال الحاكمي قبل التخطيط؛ ثم تستخدم MCD لقياس الفرق والتكامل والزمن والوكالة والغاية والسياق مقابل الضوضاء. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$\text{Q-MCD} = \Omega(Q) \times [\lambda 1D + \lambda 2I + \lambda 3T + \lambda 4A + \lambda 5P + \lambda 6C - \lambda 7N]$$

آلة الحالة المصغرة

Input	Goal Contract وتعمل أنظمة كثيرة على حل السؤال المعطى حتى لو كان السؤال رديئًا أو لا يغير القرار، فتستهلك الموارد داخل فضاء بحث غير حاكم.
Control	توليد أسئلة منافسة قبل التنفيذ
Transform	تعيين أوزان λ بحسب نوع المهمة
Evidence	إعادة السؤال عند فشل الدليل أو تغير العالم
Stop	بلوغ عتبة Question Yield أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت توليد أسئلة منافسة قبل التنفيذ وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Decision Difference فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 1 · الفصل 4 · الأدوات وعقود التنفيذ

M-AGIOS / 4

تكامل MCD و Q-MCD مع النواة

Integrating MCD and Q-MCD into the Kernel

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُقوّم جودة السؤال

Question-Quality Scorer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، توليد أسئلة منافسة قبل التنفيذ

المخرج: Artifact منظم يفعّل تعيين أوزان λ بحسب نوع المهمة

الحارس: لا تتجاوز خطر تقديس المعادلة.

الأداة 2 · مُركب أوزان MCD

MCD Weight Composer

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Noise Penalty

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Noise Penalty.

Runtime Contract المختصر

```
skill: question-quality-scorer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Question Yield, Decision Difference, Noise Penalty
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

تكامل MCD و Q-MCD مع النواة

Integrating MCD and Q-MCD into the Kernel

VA · Validation, Failure Modes, and Integration

حالة تطبيق في تحقيق تقني، يستبدل النظام سؤال: لماذا فشل الخادم؟ بسلسلة أسئلة قابلة للتمييز بين الشبكة والهوية والموارد والتكوين.

Protocol التجربة

1. ابن Baseline يعالج تكامل MCD و Q-MCD مع النواة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل توليد أسئلة منافسة قبل التنفيذ.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تقديس المعادلة: اختبار مدخلات مضادة.
- أوزان اعتباطية: اختبار مصدر أو أداة مستقلة.
- تحسين المؤشر بدل حل المشكلة: اختبار استرداد وحدود.

مقاييس الفصل

Question Yield	مقياس النتيجة الأساسية.
Decision Difference	مقياس جودة المسار أو الدليل.
Noise Penalty	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: AgentBench: Evaluating LLMs as Agents · Xiao Liu et al. · R18 | Tree of Thoughts: Deliberate Problem Solving with Large Language Models · Shunyu Yao et al. · R12

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون توليد أسئلة منافسة قبل التنفيذ
- من دون تعيين أوزان λ بحسب نوع المهمة
- من دون Question-Quality Scorer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 1 · الفصل 5 · الأطروحة والحدود

M-AGIOS / 5

العمومية كتحويل لا كمخزون

Generality as Transformation, Not Inventory

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل قد يملك النظام آلاف المهارات ولا يعرف متى يعيد تركيبها، فتكون الوفرة مستودعًا ساكنًا لا قدرة عامة.

الأطروحة المساعفية

العمومية هي تحويل من تمثيل مشكلة إلى آخر مع حفظ القيود والغاية؛ تقاس بقلة الأمثلة اللازمة، وبالقدرة على تفسير ما نُقل وما لم يُنقل.

الحد البنيوي الأدنى

يجب أن تظهر استخراج البنية الثابتة بين المجالات في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتحويل البنيوي

Moussaif Structural Transfer Theory

تقرر أن استخراج البنية الثابتة بين المجالات ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح العمومية كتحويل لا كمخزون. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للثابت المتنقل

Moussaif Transferable-Invariant Principle

تربط ترجمة القيود لا الألفاظ بـInvariant Retention، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم نقل الوكالة من سلوك لغوي متقطع إلى نظام محكوم بالحالة والدليل والأثر.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Structural Transfer Theory ليس دليلًا على وجود الآلية
- تنفيذ استخراج البنية الثابتة بين المجالات يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Few-Shot Adaptation
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من ترجمة القيود لا الألفاظ الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

العمومية كتحويل لا كمخزون

Generality as Transformation, Not Inventory

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
استخراج البنية الثابتة بين المجالات	ترجمة القيود لا الألفاظ	اختبار النقل العكسي لمنع التشابه السطحي

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. العمومية هي تحويل من تمثيل مشكلة إلى آخر مع حفظ القيود والغاية؛ تقاس بقلة الأمثلة اللازمة، وبالقدرة على تفسير ما نقل وما لم ينقل. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$T_{general} = I_{preserved} \times A_{adapted} \times V_{new} / (1 + N_{transfer})$$

آلة الحالة المصغرة

Input	Goal Contract وقد يملك النظام آلاف المهارات ولا يعرف متى يعيد تركيبها، فتكون الوفرة مستودعا ساكنا لا قدرة عامة.
Control	استخراج البنية الثابتة بين المجالات
Transform	ترجمة القيود لا الألفاظ
Evidence	اختبار النقل العكسي لمنع التشابه السطحي
Stop	بلوغ عتبة Few-Shot Adaptation أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت استخراج البنية الثابتة بين المجالات وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Invariant Retention فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 1 · الفصل 5 · الأدوات وعقود التنفيذ

M-AGIOS / 5

العمومية كتحويل لا كمخزون

Generality as Transformation, Not Inventory

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مستخرج الثوابت

Invariant Extractor

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، استخراج البنية الثابتة بين المجالات

المخرج: Artifact منظم يفعّل ترجمة القيود لا الألفاظ

الحارس: لا تتجاوز خطر نقل لغوي سطحي.

الأداة 2 · مختبر النقل المتقاطع

Cross-Domain Transfer Lab

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Negative Transfer Rate

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Negative Transfer Rate.

Runtime Contract المختصر

```
skill: invariant-extractor
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Few-Shot Adaptation, Invariant Retention, Negative Transfer Rate
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

العمومية كتحويل لا كمخزون

Generality as Transformation, Not Inventory

VA · Validation, Failure Modes, and Integration

حالة تطبيق ينقل النظام مفهوم Transaction من قاعدة بيانات إلى عملية إرسال بريد جماعي، لكنه يحتفظ بفرق قابلية التراجع وحدود الأثر الخارجي.

Protocol التجربة

1. ابنِ Baseline يعالج العمومية كتحويل لا كمخزون من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل استخراج البنية الثابتة بين المجالات.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- نقل لغوي سطحي: اختبار مدخلات مضادة.
- إسقاط قيود المجال: اختبار مصدر أو أداة مستقلة.
- ثقة زائدة عند التشابه: اختبار استرداد وحدود.

مقاييس الفصل

Few-Shot Adaptation	مقياس النتيجة الأساسية.
Invariant Retention	مقياس جودة المسار أو الدليل.
Negative Transfer Rate	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R01 · Alan M. Turing · Computing Machinery and Intelligence | R19 · Gr éGoire Mialon et al. · GAIA: A Benchmark for General AI Assistants

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون استخراج البنية الثابتة بين المجالات
- من دون ترجمة القيود لا الألفاظ
- من دون Invariant Extractor
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 2 · الفصل 6 · الأطروحة والحدود

M-AGIOS / 6

الدستور التنفيذي والهوية المستمرة

Executable Constitution and Persistent Identity

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تضيع الهوية والسلوك الحاكم عند ضغط السياق أو تبديل النموذج أو دخول وكيل فرعي، فتتحول المبادئ إلى تذكير هش.

الأطروحة المساعفية

الدستور التنفيذي مجموعة قواعد ذات أولوية وآليات إنفاذ واختبارات عدم رجوع، تُحمّل في كل Runtime وتُراجع عند كل فعل حساس.

الحد البنيوي الأدنى

يجب أن تظهر ترتيب القواعد حسب السلطة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لاستمرارية الهوية

Moussaif Identity-Continuity Theory

تقرر أن ترتيب القواعد حسب السلطة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الدستور التنفيذي والهوية المستمرة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لأولوية الدستور

Moussaif Constitutional Precedence

تربط اختبارات دستورية قبل وبعد الأفعال بـCross-Model Consistency، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تثبيت الهوية والغاية والصلاحيات والمخاطر بوصفها خدمات نواة لا زخارف Prompt.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Identity-Continuity Theory ليس دليلًا على وجود الآلية
- تنفيذ ترتيب القواعد حسب السلطة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Rule Enforcement Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من اختبارات دستورية قبل وبعد الأفعال الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 2 · الفصل 6 · الآلية والنموذج التشغيلي

M-AGIOS / 6

الدستور التنفيذي والهوية المستمرة

Executable Constitution and Persistent Identity

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
ترتيب القواعد حسب السلطة	اختبارات دستورية قبل وبعد الأفعال	توريث الحد الأدنى من القواعد للوكلاء الفرعيين

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم يُفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. الدستور التنفيذي مجموعة قواعد ذات أولوية وآليات إنفاذ واختبارات عدم رجوع، تُحمّل في كل Runtime وتراجع عند كل فعل حساس. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$C_{runtime} = P_{priority} \times E_{enforce} \times T_{trace} \times R_{regression}$$

آلة الحالة المصغرة

Input	Goal Contract وتضيع الهوية والسلوك الحاكم عند ضغط السياق أو تبديل النموذج أو دخول وكيل فرعي، فتتحول المبادئ إلى تذكير هش.
Control	ترتيب القواعد حسب السلطة
Transform	اختبارات دستورية قبل وبعد الأفعال
Evidence	توريث الحد الأدنى من القواعد للوكلاء الفرعيين
Stop	بلوغ عتبة Rule Enforcement Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت ترتيب القواعد حسب السلطة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Cross-Model Consistency فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 2 · الفصل 6 · الأدوات وعقود التنفيذ

M-AGIOS / 6

الدستور التنفيذي والهوية المستمرة

Executable Constitution and Persistent Identity

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُصرّف الدستور

Constitution Compiler

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، ترتيب القواعد حسب السلطة

المخرج: Artifact منظم يفعّل اختبارات دستورية قبل وبعد الأفعال

الحارس: لا تتجاوز خطر تناقض القواعد.

الأداة 2 · اختبار عدم الرجوع الدستوري

Constitution Regression Test

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Override Traceability

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Override Traceability.

Runtime Contract المختصر

```
skill: constitution-compiler
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Rule Enforcement Rate, Cross-Model Consistency, Override Traceability
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الدستور التنفيذي والهوية المستمرة

Executable Constitution and Persistent Identity

VA · Validation, Failure Modes, and Integration

حالة تطبيق بعد نقل المهمة من Hermes إلى Codex، تعاد اختبارات الإذن والخصوصية والتسمية قبل استكمال التنفيذ، ولا يعتمد النظام على الذاكرة الحوارية وحدها.

Protocol التجربة

1. ابن Baseline يعالج الدستور التنفيذي والهوية المستمرة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل ترتيب القواعد حسب السلطة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تناقض القواعد: اختبار مدخلات مضادة.
- دستور غير قابل للاختبار: اختبار مصدر أو أداة مستقلة.
- إخفاء سياسة المنتج كحقيقة أخلاقية: اختبار استرداد وحدود.

مقاييس الفصل

Rule Enforcement Rate	مقياس النتيجة الأساسية.
Cross-Model Consistency	مقياس جودة المسار أو الدليل.
Override Traceability	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R24 · Yuntao Bai et al. · Constitutional AI: Harmlessness from AI Feedback | R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون ترتيب القواعد حسب السلطة
- من دون اختبارات دستورية قبل وبعد الأفعال
- من دون Constitution Compiler
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 2 · الفصل 7 · الأطروحة والحدود

M-AGIOS / 7

مصرّف النية والهدف

Intent and Goal Compiler

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل الأوامر الطبيعية غامضة: قد تصف نتيجة مرغوبة من دون نطاق أو ممنوعات أو معيار انتهاء أو صاحب قرار.

الأطروحة المساعفية

يحوّل مُصرّف النية المساعفي الطلب إلى Goal Contract يضم المخرجات والقيود والأذونات وميزانية الخطأ وشروط التوقف وفرضيات قابلة للتأكيد.

الحد البنيوي الأدنى

يجب أن تظهر فصل Outcome عن Method في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتجميع النية

Moussaif Intent Compilation Theory

تقرر أن فصل Outcome عن Method ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح مصرّف النية والهدف. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للغموض المنتج

Moussaif Productive-Ambiguity Principle

تربط استخراج القيود الصريحة والضمنية بـValue of Clarification، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تثبيت الهوية والغاية والصلاحيات والمخاطر بوصفها خدمات نواة لا زخارف Prompt.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Intent Compilation Theory ليس دليلًا على وجود الآلية
- تنفيذ فصل Outcome عن Method يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Requirement Coverage
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من استخراج القيود الصريحة والضمنية الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القِسم 2 · الفصل 7 · الآلية والنموذج التشغيلي

M-AGIOS / 7

مصرّف النية والهدف

Intent and Goal Compiler

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل Outcome عن Method	استخراج القيود الصريحة والضمنية	تصنيف الغموض إلى مانع وغير مانع

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يحوّل مُصرّف النية المساعفي الطلب إلى Goal Contract يضم المخرجات والقيود والأذونات وميزانية الخطأ وشروط التوقف وفرضيات قابلة للتأكيد. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$G_{contract} = O_{outcome} + C_{constraints} + A_{authority} + S_{stop} - U_{ambiguity}$$

آلة الحالة المصغرة

Input	Goal Contract والأوامر الطبيعية غامضة: قد تصف نتيجة مرغوبة من دون نطاق أو ممنوعات أو معيار انتهاء أو صاحب قرار.
Control	فصل Outcome عن Method
Transform	استخراج القيود الصريحة والضمنية
Evidence	تصنيف الغموض إلى مانع وغير مانع
Stop	بلوغ عتبة Requirement Coverage أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل Outcome عن Method وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسّن Outcome وانهار Clarification Value فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القِسم 2 · الفصل 7 · الأدوات وعقود التنفيذ

M-AGIOS / 7

مصرّف النية والهدف

Intent and Goal Compiler

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُصرّف النية المساعفي

Moussaif Intent Compiler

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل Outcome عن Method

المخرج: Artifact منظم يفعّل استخراج القيود الصريحة والضمنية

الحارس: لا تتجاوز خطر إفراط في الاستيضاح.

الأداة 2 · بطاقة عقد الهدف

Goal-Contract Card

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Goal Drift

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Goal Drift.

Runtime Contract المختصر

```
skill: moussaif-intent-compiler
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Requirement Coverage, Clarification Value, Goal Drift
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

مصرّف النية والهدف
Intent and Goal Compiler

VA · Validation, Failure Modes, and Integration

حالة تطبيق طلب إنشاء كتاب يتحول إلى عقد يحدد 250 صفحة فعلية، العربية RTL، خط Amiri، بنية بحثية، نسبة المصطلحات الإنجليزية، والتحقق البصري.

Protocol التجربة

1. ابنِ Baseline يعالج مُصرّف النية والهدف من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل فصل Outcome عن Method.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- إفراط في الاستيضاح: اختبار مدخلات مضادة.
- افتراض يغير الغاية: اختبار مصدر أو أداة مستقلة.
- خلط الوسيلة بالنتيجة: اختبار استرداد وحدود.

مقاييس الفصل

Requirement Coverage	مقياس النتيجة الأساسية.
Clarification Value	مقياس جودة المسار أو الدليل.
Goal Drift	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R02 · Allen Newell and Herbert A. Simon | Human Problem Solving · R03 · Michael Bratman · Intention, Plans and Practical Reason

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل Outcome عن Method
- من دون استخراج القيود الصريحة والضمنية
- من دون Moussaif Intent Compiler
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 2 · الفصل 8 · الأطروحة والحدود

M-AGIOS / 8

نموذج العالم وسجل الحالة
World Model and State Ledger

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل ينفذ الوكيل بناء على لقطة قديمة أو سياق ناقص، ولا يميز بين ما كان صحيحًا وما لا يزال صحيحًا الآن.

الأطروحة المساعفية

يفصل M-AGIOS بين World State المرصود، وBelief State المستنتج، وDesired State المستهدف؛ وكل انتقال يحمل حدثًا ودليلًا وزمنًا.

الحد البنيوي الأدنى

يجب أن تظهر سجل أحداث Append-Only في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لثلاثية الحالة

Moussaif Three-State Theory

تقرر أن سجل أحداث Append-Only ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح نموذج العالم وسجل الحالة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للأثر المؤرخ

Moussaif Timestamped-Effect Law

تربط مقارنة الحالة قبل الفعل وبعده بـTransition Validity، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تثبيت الهوية والغاية والصلاحيات والمخاطر بوصفها خدمات نواة لا زخارف Prompt.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Three-State Theory ليس دليلًا على وجود الآلية
- تنفيذ سجل أحداث Append-Only يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض State Freshness
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من مقارنة الحالة قبل الفعل وبعده الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 2 · الفصل 8 · الآلية والنموذج التشغيلي

M-AGIOS / 8

نموذج العالم وسجل الحالة
World Model and State Ledger

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
سجل أحداث Append-Only	مقارنة الحالة قبل الفعل وبعده	تحديث احتمالي عند الملاحظة المتعارضة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يفصل M-AGIOS بين World State المرصود، وBelief State المستنتج، وDesired State المستهدف؛ وكل انتقال يحمل حدثا ودليلا وزمنا. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$B_t = Update(B_{(t-1)}, O_t, Q_{source}, \Delta t)$$

آلة الحالة المصغرة

Input	Goal Contract وينفذ الوكيل بناء على لقطة قديمة أو سياق ناقص، ولا يميز بين ما كان صحيحا وما لا يزال صحيحا الآن.
Control	سجل أحداث Append-Only
Transform	مقارنة الحالة قبل الفعل وبعده
Evidence	تحديث احتمالي عند الملاحظة المتعارضة
Stop	بلوغ عتبة State Freshness أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت سجل أحداث Append-Only وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Transition Validity فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 2 · الفصل 8 · الأدوات وعقود التنفيذ

M-AGIOS / 8

نموذج العالم وسجل الحالة
World Model and State Ledger

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · دفتر حالة العالم
World-State Ledger

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، سجل أحداث Append-Only

المخرج: Artifact منظم يفعّل مقارنة الحالة قبل الفعل وبعده

الحارس: لا تتجاوز خطر خلط الاعتقاد بالملاحظة.

الأداة 2 · كاشف انجراف الحالة
State-Drift Detector

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Drift Detection Time

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Drift Detection Time.

Runtime Contract المختصر

```
skill: world-state-ledger
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: State Freshness, Transition Validity, Drift Detection Time
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل
- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل
- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events
tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ
خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 2 · الفصل 8 · التحقق والفشل والتكامل

M-AGIOS / 8

نموذج العالم وسجل الحالة
World Model and State Ledger

VA · Validation, Failure Modes, and Integration

حالة تطبيق لا يقرر النظام أن موقعًا منشور لمجرد نجاح أمر Deploy؛ ينتظر Observation مستقلة من الرابط وحالة الاستضافة ثم يسجل الانتقال.

Protocol التجربة

1. ابنِ Baseline يعالج نموذج العالم وسجل الحالة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل سجل أحداث Append-Only.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- خلط الاعتقاد بالملاحظة: اختبار مدخلات مضادة.
- تحديث مدمر للتاريخ: اختبار مصدر أو أداة مستقلة.
- قراءة لاحقة تنسب إلى زمن سابق: اختبار استرداد وحدود.

مقاييس الفصل

State Freshness	مقياس النتيجة الأساسية.
Transition Validity	مقياس جودة المسار أو الدليل.
Drift Detection Time	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R13 · Joon Sung Park et al. · Generative Agents: Interactive Simulacra of Human Behavior | R17 · Shuyan Zhou et al. · WebArena: A Realistic Web Environment for Building Autonomous Agents

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون سجل أحداث Append-Only
- من دون مقارنة الحالة قبل الفعل وبعده
- من دون World-State Ledger
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 2 · الفصل 9 · الأطروحة والحدود

M-AGIOS / 9

آلة الحالة الوكيلية ودورة الحياة

Agentic State Machine and Lifecycle

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تتحول الحلقات الوكيلية إلى تدفق نصي يصعب إيقافه أو استئنافه أو تفسير موقعه من المهمة.

الأطروحة المساعفية

تُمثل المهمة كآلة حالات صريحة: Intake، Contracted، Planned، Acting، Observing، Verifying، Committed، Rolled Back أو Escalated.

الحد البنيوي الأدنى

يجب أن تظهر انتقالات بحراس Guard Conditions في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لدورة الحياة الوكيلية

Moussaif Agent Lifecycle Theory

تقرر أن انتقالات بحراس Guard Conditions ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح آلة الحالة الوكيلية ودورة الحياة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للالتزام المتحقق

Moussaif Verified-Commit Principle

تربط حفظ Checkpoint عند حدود الأثر بـResume Fidelity، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تثبيت الهوية والغاية والصلاحيات والمخاطر بوصفها خدمات نواة لا زخارف Prompt.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Agent Lifecycle Theory ليس دليلًا على وجود الآلية
- تنفيذ انتقالات بحراس Guard Conditions يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Illegal Transition Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من حفظ Checkpoint عند حدود الأثر الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

آلة الحالة الوكيلية ودورة الحياة

Agentic State Machine and Lifecycle

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
انتقالات بحراس Guard Conditions	حفظ Checkpoint عند حدود الأثر	منع الانتقال إلى Committed بلا تحقق

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تُمثل المهمة كآلة حالات صريحة: Intake، Contracted، Planned، Acting، Observing، Verifying، Committed، Rolled Back أو Escalated. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$S_{t+1} = \delta(S_t, Event_t) \text{ iff } Guard(S_t, Event_t)=true$$

آلة الحالة المصغرة

Input	Goal Contract وتحول الحلقات الوكيلية إلى تدفق نصي يصعب إيقافه أو استئنافه أو تفسير موقعه من المهمة.
Control	انتقالات بحراس Guard Conditions
Transform	حفظ Checkpoint عند حدود الأثر
Evidence	منع الانتقال إلى Committed بلا تحقق
Stop	بلوغ عتبة Illegal Transition Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت انتقالات بحراس Guard Conditions وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Resume Fidelity فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 2 · الفصل 9 · الأدوات وعقود التنفيذ

M-AGIOS / 9

آلة الحالة الوكيلية ودورة الحياة

Agentic State Machine and Lifecycle

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مصور آلة الحالة

Agent-State Visualizer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، انتقالات بحراس Guard Conditions

المخرج: Artifact منظم يفعّل حفظ Checkpoint عند حدود الأثر

الحارس: لا تتجاوز خطر حالة خفية في المحادثة.

الأداة 2 · حارس الانتقال

Transition Guard

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Checkpoint Coverage

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Checkpoint Coverage.

Runtime Contract المختصر

```
skill: agent-state-visualizer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Illegal Transition Rate, Resume Fidelity, Checkpoint Coverage
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

آلة الحالة الوكيلية ودورة الحياة
Agentic State Machine and Lifecycle

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا انقطع التنفيذ بعد رفع ملف وقبل تسجيل هويته، يستأنف النظام من Checkpoint يفحص الرفع أولًا بدل إعادة الرفع عشوائيًا.

Protocol التجربة

1. ابنِ Baseline يعالج آلة الحالة الوكيلية ودورة الحياة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل انتقالات بحراس Guard Conditions.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- حالة خفية في المحادثة: اختبار مدخلات مضادة.
- تكرار الفعل بعد الاستئناف: اختبار مصدر أو أداة مستقلة.
- Commit مبكر: اختبار استرداد وحدود.

مقاييس الفصل

Illegal Transition Rate	مقياس النتيجة الأساسية.
Resume Fidelity	مقياس جودة المسار أو الدليل.
Checkpoint Coverage	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R09 · Shunyu Yao et al. · ReAct: Synergizing Reasoning and Acting in Language Models | R15 · Charles Packer et al. · MemGPT: Towards LLMs as Operating Systems

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون انتقالات بحراس Guard Conditions
- من دون حفظ Checkpoint عند حدود الأثر
- من دون Agent-State Visualizer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 2 · الفصل 10 · الأطروحة والحدود

M-AGIOS / 10

سياسة المخاطر وميزانية الأثر

Risk Policy and Effect Budget

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل قد تكون المهمة بسيطة لغويًا لكنها عالية الأثر؛ إرسال رسالة أو حذف ملف أخطر من تلخيص طويل وإن احتاج Tokens أقل.

الأطروحة المساعفية

تخصص النواة Effect Budget مستقلًا عن Compute Budget، وتربط الاستقلالية بحاصل عدم اليقين وحجم الأثر وعدم قابلية العكس وحساسية البيانات.

الحد البنيوي الأدنى

يجب أن تظهر تصنيف الأفعال قبل التنفيذ في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لميزانية الأثر

Moussaif Effect-Budget Theory

تقرر أن تصنيف الأفعال قبل التنفيذ ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح سياسة المخاطر وميزانية الأثر. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للتصعيد المتناسب

Moussaif Proportional-Escalation Law

تربط رفع عتبة الإذن مع المخاطرة بـReversibility Score، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تثبيت الهوية والغاية والصلاحيات والمخاطر بوصفها خدمات نواة لا زخارف Prompt.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Effect-Budget Theory ليس دليلًا على وجود الآلية
- تنفيذ تصنيف الأفعال قبل التنفيذ يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Unauthorized Effect Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من رفع عتبة الإذن مع المخاطرة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

M-AGIOS / 10

سياسة المخاطر وميزانية الأثر

Risk Policy and Effect Budget

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تصنيف الأفعال قبل التنفيذ	رفع عتبة الإذن مع المخاطرة	اختيار خطوات تمهيدية قابلة للعكس

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تخصص النواة Effect Budget مستقلًا عن Compute Budget، وتربط الاستقلالية بحاصل عدم اليقين وحجم الأثر وعدم قابلية العكس وحساسية البيانات. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

$$R_{effect} = U \times I \times (1 - Rev) \times D_{sensitivity}$$

آلة الحالة المصغرة

Input	Goal Contract وقد تكون المهمة بسيطة لغويًا لكنها عالية الأثر
Control	تصنيف الأفعال قبل التنفيذ
Transform	رفع عتبة الإذن مع المخاطرة
Evidence	اختيار خطوات تمهيدية قابلة للعكس
Stop	بلوغ عتبة Unauthorized Effect Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تصنيف الأفعال قبل التنفيذ وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Reversibility Score فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 2 · الفصل 10 · الأدوات وعقود التنفيذ

M-AGIOS / 10

سياسة المخاطر وميزانية الأثر

Risk Policy and Effect Budget

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · حاسبة ميزانية الأثر

Effect-Budget Calculator

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تصنيف الأفعال قبل التنفيذ

المخرج: Artifact منظم يفعّل رفع عتبة الإذن مع المخاطرة

الحارس: لا تتجاوز خطر تقدير خطر لغوي.

الأداة 2 · بوابة التصعيد

Escalation Gate

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Escalation Precision

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Escalation Precision.

Runtime Contract المختصر

```
skill: effect-budget-calculator
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Unauthorized Effect Rate, Reversibility Score, Escalation Precision
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

سياسة المخاطر وميزانية الأثر

Risk Policy and Effect Budget

VA · Validation, Failure Modes, and Integration

حالة تطبيق يمكن للوكيل مسح سجل بريد للقراءة، لكنه لا يرسل إلى 300 جهة لأن حجم الأثر الخارجي يحتاج موافقة وهوية مستلمين وحدود امتثال.

Protocol التجربة

1. ابنِ Baseline يعالج سياسة المخاطر وميزانية الأثر من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تصنيف الأفعال قبل التنفيذ.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تقدير خطر لغوي: اختبار مدخلات مضادة.
- إذن واسع دائم: اختبار مصدر أو أداة مستقلة.
- تجاهل الأثر التراكمي: اختبار استرداد وحدود.

مقاييس الفصل

Unauthorized Effect Rate	مقياس النتيجة الأساسية.
Reversibility Score	مقياس جودة المسار أو الدليل.
Escalation Precision	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: NIST · AI RMF 1.0) Artificial Intelligence Risk Management Framework · R25 | Model Context Protocol · MCP Specification 25-11-2025 · R26

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تصنيف الأفعال قبل التنفيذ
- من دون رفع عتبة الإذن مع المخاطرة
- من دون Effect-Budget Calculator
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 3 · الفصل 11 · الأطروحة والحدود

M-AGIOS / 11

تصنيف المهمة قبل الاستدلال

Task Typing Before Inference

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُرسل كل طلب إلى النمط نفسه من التفكير، مع أن البحث والتلخيص والتنفيذ والتصميم والقرار الحساس تحتاج عقودًا وأدوات مختلفة.

الأطروحة المساعفية

تبدأ الجدولة بمُصنّف مهمة متعدد الملصقات يحدد نمط المعرفة، وحاجة الحداثة، ونوع الأثر، وطول الأفق، ودرجة الحتمية المطلوبة.

الحد البنيوي الأدنى

يجب أن تظهر استخراج سمات المهمة لا موضوعها فقط في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لنمط المهمة

Moussaif Task-Mode Theory

تقرر أن استخراج سمات المهمة لا موضوعها فقط ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تصنيف المهمة قبل الاستدلال. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لإعادة التصنيف الحدثي

Moussaif Eventful-Retyping Law

تربط تعيين Runtime Profile لكل نوع بـ Mode-Regret، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم اختيار النموذج والمهارة والمسار والميزانية على أساس المنفعة المتوقعة والدليل.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Task-Mode Theory ليس دليلًا على وجود الآلية
- تنفيذ استخراج سمات المهمة لا موضوعها فقط يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Routing Accuracy
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تعيين Runtime Profile لكل نوع الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 3 · الفصل 11 · الآلية والنموذج التشغيلي

M-AGIOS / 11

تصنيف المهمة قبل الاستدلال

Task Typing Before Inference

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
استخراج سمات المهمة لا موضوعها فقط	تعيين Runtime Profile لكل نوع	إعادة التصنيف عند ظهور دليل جديد

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تبدأ الجدولة بمُصنّف مهمة متعدد الملصقات يحدد نمط المعرفة، وحاجة الحداثة، ونوع الأثر، وطول الأفق، ودرجة الحتمية المطلوبة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$\text{Type}(t) = \text{argmax}_k \, P(k \mid \text{goal}, \text{evidence}, \text{effect}, \text{horizon})$$

آلة الحالة المصغرة

Input	Goal Contract ويُرسل كل طلب إلى النمط نفسه من التفكير، مع أن البحث والتلخيص والتنفيذ والتصميم والقرار الحساس تحتاج عقودا وأدوات مختلفة.
Control	استخراج سمات المهمة لا موضوعها فقط
Transform	تعيين Runtime Profile لكل نوع
Evidence	إعادة التصنيف عند ظهور دليل جديد
Stop	بلوغ عتبة Routing Accuracy أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت استخراج سمات المهمة لا موضوعها فقط وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Mode-Regret فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 3 · الفصل 11 · الأدوات وعقود التنفيذ

M-AGIOS / 11

تصنيف المهمة قبل الاستدلال

Task Typing Before Inference

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُصنّف المهمة الوعائي

Vascular Task Classifier

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، استخراج سمات المهمة لا موضوعها فقط

المخرج: Artifact منظم يفعّل تعيين Runtime Profile لكل نوع

الحارس: لا تتجاوز خطر تصنيف أحادي.

الأداة 2 · مولّد ملف التشغيل

Runtime-Profile Generator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Retyping Latency

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Retyping Latency.

Runtime Contract المختصر

```
skill: vascular-task-classifier
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Routing Accuracy, Mode-Regret, Retyping Latency
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

تصنيف المهمة قبل الاستدلال

Task Typing Before Inference

VA · Validation, Failure Modes, and Integration

حالة تطبيق طلب مراجعة كود يبقى Read-Only حتى يطلب المستخدم الإصلاح؛ عندها ينتقل من Diagnostic إلى Mutating Runtime Profile بصلاحيات واختبارات جديدة.

Protocol التجربة

1. ابنِ Baseline يعالج تصنيف المهمة قبل الاستدلال من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل استخراج سمات المهمة لا موضوعها فقط.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تصنيف أحادي: اختبار مدخلات مضادة.
- انحياز الكلمات المفتاحية: اختبار مصدر أو أداة مستقلة.
- ثبات النوع بعد تغير المهمة: اختبار استرداد وحدود.

مقاييس الفصل

Routing Accuracy	مقياس النتيجة الأساسية.
Mode-Regret	مقياس جودة المسار أو الدليل.
Retyping Latency	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R18 · Xiao Liu et al. · AgentBench: Evaluating LLMs as Agents | R19 · Grégoire Mialon et al. · GAIA: A Benchmark for General AI Assistants

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون استخراج سمات المهمة لا موضوعها فقط
- من دون تعيين Runtime Profile لكل نوع
- من دون Vascular Task Classifier
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

M-AGIOS / 12

توجيه النماذج متعدد المستويات

Multi-Tier Model Routing

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تشغيل النموذج الأغلى على كل خطوة يهدر المال والحدود، وتشغيل نموذج صغير دائمًا يرفع أخطاء التخطيط والتحقق.

الأطروحة المساعفية

يختار Moussaif HyperRouting نموذجًا بحسب تعقيد الخطوة وحساسية الخطأ وطول السياق وأدواته وكلفة الاستدعاء، مع تصعيد قائم على دليل فشل.

الحد البنيوي الأدنى

يجب أن تظهر فصل Model Tier عن Agent Role في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتوجيه الطبقي

Moussaif Tiered-Model Routing

تقرر أن فصل Model Tier عن Agent Role ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توجيه النماذج متعدد المستويات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للتصعيد بالدليل

Moussaif Evidence-Based Escalation

تربط تقدير المنفعة الحدية للاستدعاء بـEscalation Yield، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم اختيار النموذج والمهارة والمسار والميزانية على أساس المنفعة المتوقعة والدليل.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Tiered-Model Routing ليس دليلًا على وجود الآلية
- تنفيذ فصل Model Tier عن Agent Role يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Quality per Dollar
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تقدير المنفعة الحدية للاستدعاء الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 3 · الفصل 12 · الآلية والنموذج التشغيلي

M-AGIOS / 12

توجيه النماذج متعدد المستويات

Multi-Tier Model Routing

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل Model Tier عن Agent Role	تقدير المنفعة الحدية للاستدعاء	Fallback يحفظ الحالة ولا يعيد المهمة من الصفر

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يختار Moussaif HyperRouting نموذجا بحسب تعقيد الخطوة وحساسية الخطأ وطول السياق وأدواته وكلفة الاستدعاء، مع تصعيد قائم على دليل فشل. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Route* = argmax_m [E(Quality_m) - Cost_m - Risk_m - Latency_m]

آلة الحالة المصغرة

Input	Goal Contract وتشغيل النموذج الأغلى على كل خطوة يهدر المال والحدود، وتشغيل نموذج صغير دائمًا يرفع أخطاء التخطيط والتحقق.
Control	فصل Model Tier عن Agent Role
Transform	تقدير المنفعة الحدية للاستدعاء
Evidence	Fallback يحفظ الحالة ولا يعيد المهمة من الصفر
Stop	بلوغ عتبة Quality per Dollar أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل Model Tier عن Agent Role وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Escalation Yield فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

توجيه النماذج متعدد المستويات
Multi-Tier Model Routing

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · موجّه النماذج الفائق
Moussaif HyperRouter

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل Model Tier عن Agent Role

المخرج: Artifact منظم يفعّل تقدير المنفعة الحدية للاستدعاء

الحارس: لا تتجاوز خطر Router يثق بتقييم النموذج لنفسه.

الأداة 2 · حاسبة ندم النموذج
Model-Regret Calculator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Fallback Continuity

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Fallback Continuity.

Runtime Contract المختصر

```
skill: moussaif-hyperrouter
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Quality per Dollar, Escalation Yield, Fallback Continuity
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل
- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل
- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events
tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ
خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 3 · الفصل 12 · التحقق والفشل والتكامل

M-AGIOS / 12

توجيه النماذج متعدد المستويات

Multi-Tier Model Routing

VA · Validation, Failure Modes, and Integration

حالة تطبيق يُسند استخراج Metadata إلى نموذج صغير، ويُصعّد تركيب فرضية علمية متعارضة إلى نموذج استدلال قوي، ثم يعود التحقق الحسابي إلى أداة حتمية.

Protocol التجربة

1. ابنِ Baseline يعالج توجيه النماذج متعدد المستويات من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل فصل Model Tier عن Agent Role.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Router يثق بتقييم النموذج لنفسه: اختبار مدخلات مضادة.
- تكرار الاستدعاءات: اختبار مصدر أو أداة مستقلة.
- اختلاف Schemas بين المزودين: اختبار استرداد وحدود.

مقاييس الفصل

Quality per Dollar	مقياس النتيجة الأساسية.
Escalation Yield	مقياس جودة المسار أو الدليل.
Fallback Continuity	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R10 · Timo Schick et al. · Toolformer: Language Models Can Teach Themselves to Use Tools | R16 · Qingyun Wu et al. · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل Model Tier عن Agent Role
- من دون تقدير المنفعة الحدية للاستدعاء
- من دون Moussaif HyperRouter
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

توجيه المهارات والرسم الفائق

Skill Routing and the HyperSkill Graph

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تتكدس ملفات المهارات من دون فهرسة سببية؛ فيسترجع النظام نصًا مشابهًا لغويًا لكنه غير متوافق مع الأدوات والصلاحيات والسياق.

الأطروحة المساعفية

يمثل HyperSkill Graph كل مهارة بعقود Trigger وRequires وPreconditions وVerification وFailure Modes وسجل نجاح، ويحسب توافقها مع الحالة الحالية.

الحد البنيوي الأدنى

يجب أن تظهر استرجاع مرشح ثم إعادة ترتيب سببي في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتوافق المهاري

Moussaif Skill-Compatibility Theory

تقرر أن استرجاع مرشح ثم إعادة ترتيب سببي ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توجيه المهارات والرسم الفائق. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لأقصر تركيب موثوق

Moussaif Shortest-Reliable-Composition Law

تربط حساب التعارض بين المهارات بـComposition Depth، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم اختيار النموذج والمهارة والمسار والميزانية على أساس المنفعة المتوقعة والدليل.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Skill-Compatibility Theory ليس دليلًا على وجود الآلية
- تنفيذ استرجاع مرشح ثم إعادة ترتيب سببي يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Skill Hit Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من حساب التعارض بين المهارات الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 3 · الفصل 13 · الآلية والنموذج التشغيلي

M-AGIOS / 13

توجيه المهارات والرسم الفائق

Skill Routing and the HyperSkill Graph

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
استرجاع مرشح ثم إعادة ترتيب سببي	حساب التعارض بين المهارات	تفضيل التركيب القصير القابل للاختبار

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يمثل HyperSkill Graph كل مهارة بعقود Trigger وRequires وPreconditions وVerification وFailure Modes وسجل نجاح، ويحسب توافقها مع الحالة الحالية. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Score(s) = Rel + Utility + Compat + Evidence + History - Cost - Risk - Context

آلة الحالة المصغرة

Input	Goal Contract وتكدس ملفات المهارات من دون فهرسة سببية
Control	استرجاع مرشح ثم إعادة ترتيب سببي
Transform	حساب التعارض بين المهارات
Evidence	تفضيل التركيب القصير القابل للاختبار
Stop	بلوغ عتبة Skill Hit Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت استرجاع مرشح ثم إعادة ترتيب سببي وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Composition Depth فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 3 · الفصل 13 · الأدوات وعقود التنفيذ

M-AGIOS / 13

توجيه المهارات والرسم الفائق

Skill Routing and the HyperSkill Graph

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · موجّه HyperSkill

HyperSkill Router

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، استرجاع مرشح ثم إعادة ترتيب سببي

المخرج: Artifact منظم يفعّل حساب التعارض بين المهارات

الحارس: لا تتجاوز خطر تشابه Embedding فقط.

الأداة 2 · محلل تعارض المهارات

Skill-Conflict Analyzer

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Historical Success Calibration

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Historical Success Calibration.

Runtime Contract المختصر

```
skill: hyperskill-router
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Skill Hit Rate, Composition Depth, Historical Success Calibration
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

توجيه المهارات والرسم الفائق
Skill Routing and the HyperSkill Graph

VA · Validation, Failure Modes, and Integration

حالة تطبيق عند ترحيل نطاق موقع، يركب النظام مهارات الجرد والاستبدال والاختبار والنشر، ولا يستدعي مهارة النشر قبل إثبات خلو الملفات من الرابط القديم.

Protocol التجربة

1. ابنِ Baseline يعالج توجيه المهارات والرسم الفائق من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل استرجاع مرشح ثم إعادة ترتيب سببي.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تشابه Embedding فقط: اختبار مدخلات مضادة.
- مهارة بلا تحقق: اختبار مصدر أو أداة مستقلة.
- توريث صلاحيات زائدة: اختبار استرداد وحدود.

مقاييس الفصل

Skill Hit Rate	مقياس النتيجة الأساسية.
Composition Depth	مقياس جودة المسار أو الدليل.
Historical Success Calibration	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R10 · Toolformer: Language Models Can Teach Themselves to Use Tools · Timo Schick et al. · 25-11-2025 | MCP Specification · Model Context Protocol · R26

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون استرجاع مرشح ثم إعادة ترتيب سببي
- من دون حساب التعارض بين المهارات
- من دون HyperSkill Router
- من دون Falsification Court

ثبّت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 3 · الفصل 14 · الأطروحة والحدود

M-AGIOS / 14

جدولة الميزانية والوقت

Budget and Deadline Scheduling

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تستهلك الحلقات مواردها في أول فروع البحث ثم تعجز عن التحقق أو الإخراج؛ كما لا تميز بين Deadline حقيقي ورغبة عامة في السرعة.

الأطروحة المساعفية

يقسم المجدول ميزانية Tokens وCalls ووقت الحائط ومال المزود على مراحل السؤال والتخطيط والفعل والتحقق والاحتياط، ويعيد الموازنة بالأحداث.

الحد البنيوي الأدنى

يجب أن تظهر حجز Verification Reserve في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للاحتياط التحققي

Moussaif Verification-Reserve Theory

تقرر أن حجز Verification Reserve ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح جدولة الميزانية والوقت. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للعائد الحدي للحوسبة

Moussaif Marginal-Compute Return

تربط تقدير Critical Path بـ Verification Reserve Use، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم اختيار النموذج والمهارة والمسار والميزانية على أساس المنفعة المتوقعة والدليل.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Verification-Reserve Theory ليس دليلًا على وجود الآلية
- تنفيذ حجز Verification Reserve يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Budget Overrun
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تقدير Critical Path الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

جدولة الميزانية والوقت

Budget and Deadline Scheduling

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
حجز Verification Reserve	تقدير Critical Path	إيقاف الفروع منخفضة العائد مبكرًا

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يقسم المجدول ميزانية Tokens وCalls ووقت الحائط ومال المزود على مراحل السؤال والتخطيط والفعل والتحقق والاحتياط، ويعيد الموازنة بالأحداث. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

$$B_{total} = B_{question} + B_{plan} + B_{act} + B_{verify} + B_{reserve}$$

آلة الحالة المصغرة

Input	Goal Contract وتستهلك الحلقات مواردها في أول فروع البحث ثم تعجز عن التحقق أو الإخراج
Control	حجز Verification Reserve
Transform	تقدير Critical Path
Evidence	إيقاف الفروع منخفضة العائد مبكرًا
Stop	بلوغ عتبة Budget Overrun أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت حجز Verification Reserve وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Verification Reserve Use فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 3 · الفصل 14 · الأدوات وعقود التنفيذ

M-AGIOS / 14

جدولة الميزانية والوقت

Budget and Deadline Scheduling

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مجدول الميزانية المعرفية

Cognitive Budget Scheduler

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، حجز Verification Reserve

المخرج: Artifact منظم يفعّل تقدير Critical Path

الحارس: لا تتجاوز خطر تقليل الجودة الخفي.

الأداة 2 · قاطع الفرع منخفض العائد

Low-Yield Branch Cutter

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Deadline Success

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Deadline Success.

Runtime Contract المختصر

```
skill: cognitive-budget-scheduler
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Budget Overrun, Verification Reserve Use, Deadline Success
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 3 · الفصل 14 · التحقق والفشل والتكامل

M-AGIOS / 14

جدولة الميزانية والوقت

Budget and Deadline Scheduling

VA · Validation, Failure Modes, and Integration

حالة تطبيق في كتاب طويل، تُحجز موارد للعد الصفحي واختبار الخطوط والرندر النهائي قبل توسيع فصل تجميلي لا يغير اكتمال التسليم.

Protocol التجربة

1. ابنِ Baseline يعالج جدولة الميزانية والوقت من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل حجز Verification Reserve.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تقليل الجودة الخفي: اختبار مدخلات مضادة.
- حجز احتياط مبالغ: اختبار مصدر أو أداة مستقلة.
- تجاهل كلفة الأدوات الخارجية: اختبار استرداد وحدود.

مقاييس الفصل

Budget Overrun	مقياس النتيجة الأساسية.
Verification Reserve Use	مقياس جودة المسار أو الدليل.
Deadline Success	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R12 · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments · Tianbao Xie et al. · R20 | Tree of Thoughts: Deliberate Problem Solving with Large Language Models · Shunyu Yao et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون حجز Verification Reserve
- من دون تقدير Critical Path
- من دون Cognitive Budget Scheduler
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 3 · الفصل 15 · الأطروحة والحدود

M-AGIOS / 15

المقاطعة والاستباق والتوقف

Interrupts, Preemption, and Stopping

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يواصل الوكيل خطة أصبحت باطلة بعد رسالة جديدة أو فشل أداة، أو يدخل Loop لأن معيار الانتهاء وصياغة الاستثناء غير صريحين.

الأطروحة المساعفية

تعامل النواة الأحداث الجديدة كمقاطعات ذات أولوية، تحفظ الحالة، وتقارن صلاحية الخطة، ثم تستأنف أو تعيد التخطيط أو تتوقف وفق سياسة محددة.

الحد البنيوي الأدنى

يجب أن تظهر Checkpoint قبل نقطة المقاطعة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للمقاطعة الدلالية

Moussaif Semantic-Interrupt Theory

تقرر أن Checkpoint قبل نقطة المقاطعة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح المقاطعة والاستباق والتوقف. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للتوقف العقلاني

Moussaif Rational-Stop Law

تربط تمييز Pause عن Cancel عن Override بـDuplicate-Action Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم اختيار النموذج والمهارة والمسار والميزانية على أساس المنفعة المتوقعة والدليل.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Semantic-Interrupt Theory ليس دليلًا على وجود الآلية
- تنفيذ Checkpoint قبل نقطة المقاطعة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Interrupt Response Time
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تمييز Pause عن Cancel عن Override الذي يغير النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

M-AGIOS / 15

المقاطعة والاستباق والتوقف

Interrupts, Preemption, and Stopping

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
Checkpoint قبل نقطة المقاطعة	تمييز Pause عن Override عن Cancel	تطبيق Loop Saturation على الدليل والكلفة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تعامل النواة الأحداث الجديدة كمقاطعات ذات أولوية، تحفظ الحالة، وتقارن صلاحية الخطة، ثم تستأنف أو تعيد التخطيط أو تتوقف وفق سياسة محددة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Stop iff E(ΔEvidence_next) ≤ Cost_next + Risk_next

آلة الحالة المصغرة

Input	Goal Contract ويواصل الوكيل خطة أصبحت باطلة بعد رسالة جديدة أو فشل أداة، أو يدخل Loop لأن معيار الانتهاء وصياغة الاستثناء غير صريحين.
Control	Checkpoint قبل نقطة المقاطعة
Transform	تمييز Pause عن Override عن Cancel
Evidence	تطبيق Loop Saturation على الدليل والكلفة
Stop	بلوغ عتبة Interrupt Response Time أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت Checkpoint قبل نقطة المقاطعة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Duplicate-Action Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 3 · الفصل 15 · الأدوات وعقود التنفيذ

M-AGIOS / 15

المقاطعة والاستباق والتوقف

Interrupts, Preemption, and Stopping

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · متحكم المقاطعات

Agent Interrupt Controller

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، Checkpoint قبل نقطة المقاطعة

المخرج: Artifact منظم يفعّل تمييز Pause عن Cancel عن Override

الحارس: لا تتجاوز خطر فقد الحالة.

الأداة 2 · كاشف تشبع الحلقة

Loop-Saturation Detector

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Stop-Regret

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Stop-Regret.

Runtime Contract المختصر

```
skill: agent-interrupt-controller
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Interrupt Response Time, Duplicate-Action Rate, Stop-Regret
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 3 · الفصل 15 · التحقق والفشل والتكامل

M-AGIOS / 15

المقاطعة والاستباق والتوقف

Interrupts, Preemption, and Stopping

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا عدّل المستخدم عدد الصفحات أثناء البناء، يوقف النظام التوليد عند حد آمن، يعيد حساب الخريطة، ولا يدمج النسختين في ملف متناقض.

Protocol التجربة

1. ابن Baseline يعالج المقاطعة والاستباق والتوقف من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل Checkpoint قبل نقطة المقاطعة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- فقد الحالة: اختبار مدخلات مضادة.
- اعتبار كل رسالة إلغاء: اختبار مصدر أو أداة مستقلة.
- توقف بسبب نفاد صامت: اختبار استرداد وحدود.

مقاييس الفصل

Interrupt Response Time	مقياس النتيجة الأساسية.
Duplicate-Action Rate	مقياس جودة المسار أو الدليل.
Stop-Regret	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R11 · Reflexion: Language Agents with Verbal Reinforcement Learning · Noah Shinn et al. · R15 | MemGPT: Towards LLMs as Operating Systems · Charles Packer et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون Checkpoint قبل نقطة المقاطعة
- من دون تمييز Pause عن Cancel عن Override
- من دون Agent Interrupt Controller
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 4 · الفصل 16 · الأطروحة والحدود

M-AGIOS / 16

هرمية الذاكرة المساعفية

Moussaif Memory Hierarchy

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تُعامل كل المعلومات كسياق متساوٍ، فتزاحم التفاصيل القديمة قواعد المهمة، أو تُسترجع ذاكرة بعيدة لمجرد التشابه.

الأطروحة المساعفية

تقسم الذاكرة إلى Working، Episodic، Semantic، Procedural، Project Registry وArchive؛ ولكل طبقة زمن وصول وسياسة كتابة وثقة ونسيان.

الحد البنيوي الأدنى

يجب أن تظهر عنونة الذاكرة بحسب الوظيفة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للطبقات الذاكرية

Moussaif Memory-Layer Theory

تقرر أن عنونة الذاكرة بحسب الوظيفة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح هرمية الذاكرة المساعفية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لترقية الذاكرة

Moussaif Memory-Promotion Law

تربط ترقية الخبرة من Episode إلى Procedure بعد اختبار بـPromotion Error، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تحويل السياق المحدود إلى ذاكرة هرمية قابلة للعنونة والضغط والاسترجاع والنسيان المنضبط.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Memory-Layer Theory ليس دليلًا على وجود الآلية
- تنفيذ عنونة الذاكرة بحسب الوظيفة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Retrieval Precision
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من ترقية الخبرة من Episode إلى Procedure بعد اختبار الذي يغير النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 4 · الفصل 16 · الآلية والنموذج التشغيلي

M-AGIOS / 16

هرمية الذاكرة المساعفية

Moussaif Memory Hierarchy

ME · Mechanism, State, and Equation

سلسلة التحويل

M1
عنونة الذاكرة بحسب الوظيفة

M2
ترقية الخبرة من Episode إلى Procedure بعد اختبار

M3
منع Archive من التأثير بلا إعادة تحقق

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تقسم الذاكرة إلى Working، Episodic، Semantic، Procedural، Project Registry وArchive؛ ولكل طبقة زمن وصول وسياسة كتابة وثقة ونسيان. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

M_write(layer) = f(reuse, evidence, volatility, sensitivity, cost)

آلة الحالة المصغرة

Input	Goal Contract وتُعامل كل المعلومات كسياق متساو، فتزاحم التفاصيل القديمة قواعد المهمة، أو تُسترجع ذاكرة بعيدة لمجرد التشابه.
Control	عنونة الذاكرة بحسب الوظيفة
Transform	ترقية الخبرة من Episode إلى Procedure بعد اختبار
Evidence	منع Archive من التأثير بلا إعادة تحقق
Stop	بلوغ عتبة Retrieval Precision أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت عنونة الذاكرة بحسب الوظيفة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Promotion Error فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 4 · الفصل 16 · الأدوات وعقود التنفيذ

M-AGIOS / 16

هرمية الذاكرة المساعفية

Moussaif Memory Hierarchy

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · موجّه طبقات الذاكرة

Memory-Tier Router

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، عنونة الذاكرة بحسب الوظيفة

المخرج: Artifact منظم يفعّل ترقية الخبرة من Episode إلى Procedure بعد اختبار

الحارس: لا تتجاوز خطر ذاكرة شاملة غير مفصولة.

الأداة 2 · بوابة ترقية الخبرة

Experience-Promotion Gate

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Memory Access Cost

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Memory Access Cost.

Runtime Contract المختصر

```
skill: memory-tier-router
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Retrieval Precision, Promotion Error, Memory Access Cost
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

هرمية الذاكرة المساعفية

Moussaif Memory Hierarchy

VA · Validation, Failure Modes, and Integration

حالة تطبيق يظل مسار محاولة نشر فاشلة في Episodic، ولا يتحول إلى Skill إلا بعد إثبات سبب الفشل واختبار الإجراء المصحح مرتين.

Protocol التجربة

1. ابنِ Baseline يعالج هرمية الذاكرة المساعفية من دون النظريتين المقترحتين.
2. ثبّت النموذج والـTools والميزانية، ثم فعّل عنونة الذاكرة بحسب الوظيفة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- ذاكرة شاملة غير مفصولة: اختبار مدخلات مضادة.
- ترقية تجربة فاشلة: اختبار مصدر أو أداة مستقلة.
- استرجاع قاعدة قديمة كحقيقة حالية: اختبار استرداد وحدود.

مقاييس الفصل

Retrieval Precision	مقياس النتيجة الأساسية.
Promotion Error	مقياس جودة المسار أو الدليل.
Memory Access Cost	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R15 · A-Mem: Agentic Memory for LLM Agents · Wujiang Xu et al. · R27 | MemGPT: Towards LLMs as Operating Systems · Charles Packer et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون عنونة الذاكرة بحسب الوظيفة
- من دون ترقية الخبرة من Episode إلى Procedure بعد اختبار
- من دون Memory-Tier Router
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 4 · الفصل 17 · الأطروحة والحدود

M-AGIOS / 17

افتراضية السياق والضغط القابل للفك

Context Virtualization and Reversible Compaction

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل نافذة السياق محدودة؛ والضغط الساذج يحذف الاستثناءات أو ينسب الاستنتاج إلى المستخدم أو يفقد علاقات المشروع.

الأطروحة المساعفية

يستخدم M-AGIOS صفحات Context ذات معرفات وموجزات قابلة للفك وClaim Pointers، ويحمّل أقل Working Set يحقق المهمة مع سجل ما تم إسقاطه.

الحد البنيوي الأدنى

يجب أن تظهر فصل Hot Context عن Cold Context في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للسياق الافتراضي

Moussaif Virtual-Context Theory

تقرر أن فصل Hot Context عن Cold Context ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح افتراضية السياق والضغط القابل للفك. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للضغط القابل للفك

Moussaif Reversible-Compaction Principle

تربط ضغط يحتفظ بقيود وإسناد بـCompaction Ratio، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تحويل السياق المحدود إلى ذاكرة هرمية قابلة للعنونة والضغط والاسترجاع والنسيان المنضبط.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Virtual-Context Theory ليس دليلًا على وجود الآلية
- تنفيذ فصل Hot Context عن Cold Context يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Constraint Recall
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من ضغط يحتفظ بقيود وإسناد الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 4 · الفصل 17 · الآلية والنموذج التشغيلي

M-AGIOS / 17

افتراضية السياق والضغط القابل للفك

Context Virtualization and Reversible Compaction

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل Hot Context عن Cold Context	ضغط يحتفظ بقيود وإسناد	Demand Paging عند إشارة نقص

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يستخدم M-AGIOS صفحات Context ذات معرفات وموجزات قابلة للفك وClaim Pointers، ويحمّل أقل Working Set يحقق المهمة مع سجل ما تم إسقاطه. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

C_effective = WorkingSet + Pointers + OnDemand(Page_i)

آلة الحالة المصغرة

Input	Goal Contract ونافذة السياق محدودة
Control	فصل Hot Context عن Cold Context
Transform	ضغط يحتفظ بقيود وإسناد
Evidence	Demand Paging عند إشارة نقص
Stop	بلوغ عتبة Constraint Recall أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل Hot Context عن Cold Context وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Compaction Ratio فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 4 · الفصل 17 · الأدوات وعقود التنفيذ

M-AGIOS / 17

افتراضية السياق والضغط القابل للفك

Context Virtualization and Reversible Compaction

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مجزئ صفحات السياق

Context Pager

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل Hot Context عن Cold Context

المخرج: Artifact منظم يفعّل ضغط يحتفظ بقيود وإسناد

الحارس: لا تتجاوز خطر ملخص غير قابل للتتبع.

الأداة 2 · محكمة عدم الضياع

No-Loss Compaction Court

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Page-Fault Utility

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Page-Fault Utility.

Runtime Contract المختصر

```
skill: context-pager
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Constraint Recall, Compaction Ratio, Page-Fault Utility
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

افتراضية السياق والضغط القابل للفك

Context Virtualization and Reversible Compaction

VA · Validation, Failure Modes, and Integration

حالة تطبيق يحفظ موجز المشروع الرابط القانوني الحالي وسبب تغييره ومصدره، بينما تُنقل سجلات التنفيذ التفصيلية إلى صفحة باردة قابلة للاستدعاء.

Protocol التجربة

1. ابن Baseline يعالج افتراضية السياق والضغط القابل للفك من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل فصل Hot Context عن Cold Context.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- ملخص غير قابل للتتبع: اختبار مدخلات مضادة.
- ضغط القيود النادرة: اختبار مصدر أو أداة مستقلة.
- Page Thrashing: اختبار استرداد وحدود.

مقاييس الفصل

Constraint Recall	مقياس النتيجة الأساسية.
Compaction Ratio	مقياس جودة المسار أو الدليل.
Page-Fault Utility	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R15 · Charles Packer et al. · MemGPT: Towards LLMs as Operating Systems | R27 · Wujiang Xu et al. · A-Mem: Agentic Memory for LLM Agents

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل Hot Context عن Cold Context
- من دون ضغط يحتفظ بقيود وإسناد
- من دون Context Pager
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 4 · الفصل 18 · الأطروحة والحدود

M-AGIOS / 18

قاعدة المعرفة كنظام تشغيل

Knowledge Base as Operating System

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُنظر إلى KB كموسوعة تُقرأ عند التشابه، لا كطبقة تقرر ما يجب تفعيله وكيف يُتحقق من صلاحيته في Runtime.

الأطروحة المساعفية

KB-as-OS تجمع المعرفة التصريحية والإجرائية والسياسات وسجل المشاريع في رسم موقّع، وتخرج قرارات Routing وConstraints وTests لا فقرات اقتباس فقط.

الحد البنيوي الأدنى

يجب أن تظهر تمييز Knowledge Node عن Skill Node في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للحاكمية المعرفية

Moussaif Knowledge-Governance Theory

تقرر أن تمييز Knowledge Node عن Skill Node ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح قاعدة المعرفة كنظام تشغيل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للتفعيل السياقي

Moussaif Contextual-Activation Law

تربط ربط كل قاعدة بنطاق وصلاحية بـPolicy Coverage، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تحويل السياق المحدود إلى ذاكرة هرمية قابلة للعنونة والضغط والاسترجاع والنسيان المنضبط.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Knowledge-Governance Theory ليس دليلًا على وجود الآلية
- تنفيذ تمييز Knowledge Node عن Skill Node يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Activation Precision
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من ربط كل قاعدة بنطاق وصلاحية الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 4 · الفصل 18 · الآلية والنموذج التشغيلي

M-AGIOS / 18

قاعدة المعرفة كنظام تشغيل

Knowledge Base as Operating System

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تمييز Knowledge Node عن Skill Node	ربط كل قاعدة بنطاق وصلاحية	إنتاج Execution Context صغير من رسم كبير

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصيل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. KB-as-OS تجمع المعرفة التصريحية والإجرائية والسياسات وسجل المشاريع في رسم موقع، وتخرج قرارات Routing وConstraints وTests لا فقرات اقتباس فقط. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Pack* = min Pack such that Coverage(goal, constraints, tests) ≥ τ

آلة الحالة المصغرة

Input	Goal Contract ويُنظر إلى KB كموسوعة تُقرأ عند التشابه، لا كطبقة تقرر ما يجب تفعيله وكيف يُتحقق من صلاحيته في Runtime.
Control	تمييز Knowledge Node عن Skill Node
Transform	ربط كل قاعدة بنطاق وصلاحية
Evidence	إنتاج Execution Context صغير من رسم كبير
Stop	بلوغ عتبة Activation Precision أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تمييز Knowledge Node عن Skill Node وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Policy Coverage فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 4 · الفصل 18 · الأدوات وعقود التنفيذ

M-AGIOS / 18

قاعدة المعرفة كنظام تشغيل

Knowledge Base as Operating System

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُصرّف KB-as-OS

KB-as-OS Compiler

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تمييز Knowledge Node عن Skill Node

المخرج: Artifact منظم يفعّل ربط كل قاعدة بنطاق وصلاحية

الحارس: لا تتجاوز خطر حشو Retrieval.

الأداة 2 · مُنشئ حزمة التنفيذ

Execution-Pack Builder

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Context Pack Size

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Context Pack Size.

Runtime Contract المختصر

```
skill: kb-as-os-compiler
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Activation Precision, Policy Coverage, Context Pack Size
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

قاعدة المعرفة كنظام تشغيل
Knowledge Base as Operating System

VA · Validation, Failure Modes, and Integration

حالة تطبيق بدل تحميل 50 MB من المعرفة، يُنتج النظام حزمة تضم دستور المهمة وثلاث مهارات وسجل المشروع وخمسة اختبارات لازمة فقط.

Protocol التجربة

1. ابنِ Baseline يعالج قاعدة المعرفة كنظام تشغيل من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تمييز Knowledge Node عن Skill Node.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- حشو Retrieval: اختبار مدخلات مضادة.
- قاعدة بلا Version: اختبار مصدر أو أداة مستقلة.
- تناقض Project Registry: اختبار استرداد وحدود.

مقاييس الفصل

Activation Precision	مقياس النتيجة الأساسية.
Policy Coverage	مقياس جودة المسار أو الدليل.
Context Pack Size	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R22 · Patrick Lewis et al. · Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks | R26 · Model Context Protocol · MCP Specification · 25-11-2025

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تمييز Knowledge Node عن Skill Node
- من دون ربط كل قاعدة بنطاق وصلاحية
- من دون KB-as-OS Compiler
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 4 · الفصل 19 · الأطروحة والحدود

M-AGIOS / 19

التعلم من الحلقات والخبرة

Learning from Episodes and Experience

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تعيد الوكالات الأخطاء نفسها لأن سجل التنفيذ مجرد Logs، أو تستخلص قاعدة عامة من نجاح مصادف واحد.

الأطروحة المساعفية

تحول حلقة ما بعد التنفيذ المساعفية Episode إلى Claims سببية وCounterfactuals وCandidate Skills، ثم لا تعتمدها إلا باختبار إعادة وتشابه حدودي.

الحد البنيوي الأدنى

يجب أن تظهر فصل Outcome عن Causal Credit في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتدرج الدلالي

Moussaif Semantic-Gradient Theory

تقرر أن فصل Outcome عن Causal Credit ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التعلم من الحلقات والخبرة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للتعلم المتحفظ

Moussaif Conservative-Learning Principle

تربط إنشاء تجربة مضادة بـCausal Attribution Accuracy، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تحويل السياق المحدود إلى ذاكرة هرمية قابلة للعنونة والضغط والاسترجاع والنسيان المنضبط.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Semantic-Gradient Theory ليس دليلًا على وجود الآلية
- تنفيذ فصل Outcome عن Causal Credit يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Repeat-Failure Reduction
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من إنشاء تجربة مضادة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 4 · الفصل 19 · الآلية والنموذج التشغيلي

M-AGIOS / 19

التعلم من الحلقات والخبرة

Learning from Episodes and Experience

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل Outcome عن Causal Credit	إنشاء تجربة مضادة	تحديث Skill Success History مع سياقها

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تحولِ حلقة ما بعد التنفيذ المساعفية Episode إلى Claims سببية وCounterfactuals وCandidate Skills، ثم لا تعتمدها إلا باختبار إعادة وتشابه حدودي. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Learn = Episode + Feedback + Counterfactual + Reproduction

آلة الحالة المصغرة

Input	Goal Contract وتعيد الوكالات الأخطاء نفسها لأن سجل التنفيذ مجرد Logs، أو تستخلص قاعدة عامة من نجاح مصادف واحد.
Control	فصل Outcome عن Causal Credit
Transform	إنشاء تجربة مضادة
Evidence	تحديث Skill Success History مع سياقها
Stop	بلوغ عتبة Repeat-Failure Reduction أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل Outcome عن Causal Credit وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Causal Attribution Accuracy فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 4 · الفصل 19 · الأدوات وعقود التنفيذ

M-AGIOS / 19

التعلم من الحلقات والخبرة

Learning from Episodes and Experience

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مقطر الخبرة

Episode Distiller

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل Outcome عن Causal Credit

المخرج: Artifact منظم يفعّل إنشاء تجربة مضادة

الحارس: لا تتجاوز خطر تعلم من النتيجة فقط.

الأداة 2 · مولّد الاختبار المضاد

Counterfactual Test Generator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Skill Promotion Precision

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Skill Promotion Precision.

Runtime Contract المختصر

```
skill: episode-distiller
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Repeat-Failure Reduction, Causal Attribution Accuracy, Skill Promotion Precision
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

التعلم من الحلقات والخبرة
Learning from Episodes and Experience

VA · Validation, Failure Modes, and Integration

حالة تطبيق بعد نجاح إصلاح، يغير النظام متغيراً غير سببي في Sandbox؛ إن بقي النجاح فلا ينسب الفضل إليه عند بناء المهارة.

Protocol التجربة

1. ابنِ Baseline يعالج التعلم من الحلقات والخبرة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل فصل Outcome عن Causal Credit.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تعلم من النتيجة فقط: اختبار مدخلات مضادة.
- ذاكرة انعكاس تبريرية: اختبار مصدر أو أداة مستقلة.
- تعميم من حالة واحدة: اختبار استرداد وحدود.

مقاييس الفصل

Repeat-Failure Reduction	مقياس النتيجة الأساسية.
Causal Attribution Accuracy	مقياس جودة المسار أو الدليل.
Skill Promotion Precision	مقياس الكلفة أو الفشل المضاد.

التكامل العملي
يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R11 · Voyager: An Open-Ended Embodied Agent with Large Language Models · Guanzhi Wang et al. · R14 | Reflexion: Language Agents with Verbal Reinforcement Learning · Noah Shinn et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل Outcome عن Causal Credit
- من دون إنشاء تجربة مضادة
- من دون Episode Distiller
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 4 · الفصل 20 · الأطروحة والحدود

M-AGIOS / 20

النسيان وإعادة البناء والخصوصية

Forgetting, Reconstruction, and Privacy

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل الاحتفاظ الدائم يراكم الضوضاء والحساسية والتعارض، بينما الحذف غير الموثق يمحو الأساس الذي بُني عليه قرار سابق.

الأطروحة المساعفية

يطبق Memory Metabolism درجات اضمحلال وسياسات Retention وحذفًا مشفرًا، مع حفظ بصمة قرار غير حساسة تسمح بفهم المسار دون استعادة السر.

الحد البنيوي الأدنى

يجب أن تظهر تقدير القيمة المستقبلية للذاكرة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للنسيان الوظيفي

Moussaif Functional-Forgetting Theory

تقرر أن تقدير القيمة المستقبلية للذاكرة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح النسيان وإعادة البناء والخصوصية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للأثر الأدنى

Moussaif Minimal-Memory Trace

تربط فصل البيانات عن المشتقات بـPrivacy Exposure، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تحويل السياق المحدود إلى ذاكرة هرمية قابلة للعنونة والضغط والاسترجاع والنسيان المنضبط.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Functional-Forgetting Theory ليس دليلًا على وجود الآلية
- تنفيذ تقدير القيمة المستقبلية للذاكرة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Stale Memory Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من فصل البيانات عن المشتقات الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

النسيان وإعادة البناء والخصوصية

Forgetting, Reconstruction, and Privacy

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تقدير القيمة المستقبلية للذاكرة	فصل البيانات عن المشتقات	إعادة بناء المعرفة من مصادر أصلية عند الحاجة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يطبق Memory Metabolism درجات اضمحلال وسياسات Retention وحذفًا مشفرًا، مع حفظ بصمة قرار غير حساسة تسمح بفهم المسار دون استعادة السر. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

Keep(m) iff E(FutureUtility_m) > Storage + PrivacyRisk + DriftRisk

آلة الحالة المصغرة

Input	Goal Contract والاحتفاظ الدائم يراكم الضوضاء والحساسية والتعارض، بينما الحذف غير الموثق يمحو الأساس الذي بني عليه قرار سابق.
Control	تقدير القيمة المستقبلية للذاكرة
Transform	فصل البيانات عن المشتقات
Evidence	إعادة بناء المعرفة من مصادر أصلية عند الحاجة
Stop	بلوغ عتبة Stale Memory Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تقدير القيمة المستقبلية للذاكرة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Privacy Exposure فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 4 · الفصل 20 · الأدوات وعقود التنفيذ

M-AGIOS / 20

النسيان وإعادة البناء والخصوصية

Forgetting, Reconstruction, and Privacy

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مخطط احتفاظ الذاكرة

Memory-Retention Planner

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تقدير القيمة المستقبلية للذاكرة

المخرج: Artifact منظم يفعّل فصل البيانات عن المشتقات

الحارس: لا تتجاوز خطر حذف المصدر وبقاء Embedding.

الأداة 2 · ممحاة المشتقات الحساسة

Sensitive-Derivative Eraser

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Reconstruction Fidelity

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Reconstruction Fidelity.

Runtime Contract المختصر

```
skill: memory-retention-planner
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Stale Memory Rate, Privacy Exposure, Reconstruction Fidelity
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 4 · الفصل 20 · التحقق والفشل والتكامل

M-AGIOS / 20

النسيان وإعادة البناء والخصوصية

Forgetting, Reconstruction, and Privacy

VA · Validation, Failure Modes, and Integration

حالة تطبيق عند إبطال مفتاح API، يحذف السر وكل مشتق قابل للاسترجاع، ويبقى سجل أن تكوين المزود يحتاج مفتاحًا جديدًا دون الاحتفاظ بقيمته.

Protocol التجربة

1. ابن Baseline يعالج النسيان وإعادة البناء والخصوصية من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تقدير القيمة المستقبلية للذاكرة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- حذف المصدر وبقاء Embedding: اختبار مدخلات مضادة.
- نسيان قرار نشط: اختبار مصدر أو أداة مستقلة.
- إعادة بناء من مصدر تغير: اختبار استرداد وحدود.

مقاييس الفصل

Stale Memory Rate	مقياس النتيجة الأساسية.
Privacy Exposure	مقياس جودة المسار أو الدليل.
Reconstruction Fidelity	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R25 · NIST · Artificial Intelligence Risk Management Framework (1.0 AI RMF) | R27 · Wujiang Xu et al. · A-Mem: Agentic Memory for LLM Agents

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تقدير القيمة المستقبلية للذاكرة
- من دون فصل البيانات عن المشتقات
- من دون Memory-Retention Planner
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 5 · الفصل 21 · الأطروحة والحدود

M-AGIOS / 21

عقود الأدوات ومخططات الفعل

Tool Contracts and Action Schemas

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل قد يختار النموذج أداة صحيحة باسمها لكنه يمرر مدخلات ناقصة أو يسيء تفسير نجاح HTTP بوصفه نجاح المهمة.

الأطروحة المساعفية

كل Tool في M-AGIOS يحمل Schema دلاليًا: Preconditions وAuthority وSide Effects وIdempotency وEvidence وSuccess وRollback وTimeout.

الحد البنيوي الأدنى

يجب أن تظهر تحقق مدخلات قبل الاستدعاء في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لعقد الفعل

Moussaif Action-Contract Theory

تقرر أن تحقق مدخلات قبل الاستدعاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح عقود الأدوات ومخططات الفعل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للدليل الناتج

Moussaif Resulting-Evidence Principle

تربط تمييز Transport Success عن Task Success بـ False-Success Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم جعل الفعل الخارجي عقدًا قابلاً للتحقق والعكس، لا تخمينًا لغويًا غير محكوم.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Action-Contract Theory ليس دليلًا على وجود الآلية
- تنفيذ تحقق مدخلات قبل الاستدعاء يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Schema Validity
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تمييز Transport Success عن Task Success الذي يغير النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 5 · الفصل 21 · الآلية والنموذج التشغيلي

M-AGIOS / 21

عقود الأدوات ومخططات الفعل

Tool Contracts and Action Schemas

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تحقق مدخلات قبل الاستدعاء	تمييز Transport Success عن Task Success	ربط Tool Result بتحديث State

تبدأ الآلية بإثبات أنّ المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. كل Tool في M-AGIOS يحمل Schema دلاليًا: Preconditions وAuthority وSide Effects وIdempotency وSuccess Evidence وRollback وTimeout. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

ToolSuccess = TransportOK × PostconditionOK × EvidenceBound

آلة الحالة المصغرة

Input	Goal Contract وقد يختار النموذج أداة صحيحة باسمها لكنه يمرر مدخلات ناقصة أو يسيء تفسير نجاح HTTP بوصفه نجاح المهمة.
Control	تحقق مدخلات قبل الاستدعاء
Transform	تمييز Transport Success عن Task Success
Evidence	ربط Tool Result بتحديث State
Stop	بلوغ عتبة Schema Validity أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تحقق مدخلات قبل الاستدعاء وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار False-Success Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 5 · الفصل 21 · الأدوات وعقود التنفيذ

M-AGIOS / 21

عقود الأدوات ومخططات الفعل

Tool Contracts and Action Schemas

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُصرّف عقد الأداة

Tool-Contract Compiler

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تحقق مدخلات قبل الاستدعاء

المخرج: Artifact منظم يفعّل تمييز Transport Success عن Task Success

الحارس: لا تتجاوز خطر مدخلات مُهلوسة.

الأداة 2 · مفسر أثر الاستدعاء

Call-Effect Interpreter

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Rollback Coverage

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Rollback Coverage.

Runtime Contract المختصر

```
skill: tool-contract-compiler
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Schema Validity, False-Success Rate, Rollback Coverage
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 5 · الفصل 21 · التحقق والفشل والتكامل

M-AGIOS / 21

عقود الأدوات ومخططات الفعل

Tool Contracts and Action Schemas

VA · Validation, Failure Modes, and Integration

حالة تطبيق لا يعد رفع PDF ناجحًا حتى تعيد الأداة هوية ثابتة وحجمًا مطابقًا ويمكن للنظام فتح النسخة المحفوظة أو إثبات نتيجة النقل.

Protocol التجربة

1. ابن Baseline يعالج عقود الأدوات ومخططات الفعل من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تحقق مدخلات قبل الاستدعاء.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- مدخلات مُهلوسة: اختبار مدخلات مضادة.
- نجاح شكلي: اختبار مصدر أو أداة مستقلة.
- أداة غير Idempotent يعاد استدعاؤها: اختبار استرداد وحدود.

مقاييس الفصل

Schema Validity	مقياس النتيجة الأساسية.
False-Success Rate	مقياس جودة المسار أو الدليل.
Rollback Coverage	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R10 · Toolformer: Language Models Can Teach Themselves to Use Tools · Timo Schick et al. · 25-11-2025 | R26 · Model Context Protocol · MCP Specification

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تحقق مدخلات قبل الاستدعاء
- من دون تمييز Transport Success عن Task Success
- من دون Tool-Contract Compiler
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 5 · الفصل 22 · الأطروحة والحدود

M-AGIOS / 22

واجهات MCP وطبقة التوافق

MCP Interfaces and Compatibility Layer

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل ترتبط الوكالات بكل خدمة بتكامل خاص؛ فيصعب نقلها وتوحيد الصلاحيات والتسجيل، وتختلف دلالات الموارد والأدوات بين المزودين.

الأطروحة المساعفية

تستخدم طبقة توافق مساعفية MCP وAdapters لتوحيد Discovery وSchema وAuthorization وError Mapping، مع إبقاء دلالة الأثر في عقد محلي أعلى.

الحد البنيوي الأدنى

يجب أن تظهر اكتشاف قدرات وقت التشغيل في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لقابلية نقل الأداة

Moussaif Tool-Portability Theory

تقرر أن اكتشاف قدرات وقت التشغيل ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح واجهات MCP وطبقة التوافق. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للعقد الأعلى

Moussaif Higher-Contract Law

تربط تطبيع الأخطاء والمهل بـVersion Break Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم جعل الفعل الخارجي عقدًا قابلاً للتحقق والعكس، لا تخمينًا لغويًا غير محكوم.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Tool-Portability Theory ليس دليلًا على وجود الآلية
- تنفيذ اكتشاف قدرات وقت التشغيل يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Adapter Coverage
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تطبيع الأخطاء والمهل الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

واجهات MCP وطبقة التوافق

MCP Interfaces and Compatibility Layer

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
اكتشاف قدرات وقت التشغيل	تطبيع الأخطاء والمهل	Version Pinning واختبار توافق

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تستخدم طبقة توافق مساعفية MCP Adapters لتوحيد Discovery وSchema وAuthorization وError Mapping، مع إبقاء دلالة الأثر في عقد محلي أعلى. لا تُعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

$$\text{Compat} = \text{SchemaMatch} \times \text{SemanticMap} \times \text{AuthFit} \times \text{VersionTest}$$

آلة الحالة المصغرة

Input	Goal Contract وترتبط الوكالات بكل خدمة بتكامل خاص
Control	اكتشاف قدرات وقت التشغيل
Transform	تطبيع الأخطاء والمهل
Evidence	Version Pinning واختبار توافق
Stop	بلوغ عتبة Adapter Coverage أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت اكتشاف قدرات وقت التشغيل وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Version Break Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 5 · الفصل 22 · الأدوات وعقود التنفيذ

M-AGIOS / 22

واجهات MCP وطبقة التوافق

MCP Interfaces and Compatibility Layer

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · جسر MCP المساعفي

Moussaif MCP Bridge

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، اكتشاف قدرات وقت التشغيل

المخرج: Artifact منظم يفعّل تطبيع الأخطاء والمهل

الحارس: لا تتجاوز خطر اعتبار Schema معنى كاملا.

الأداة 2 · مصفوفة توافق الأدوات

Tool-Compatibility Matrix

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Semantic Consistency

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Semantic Consistency.

Runtime Contract المختصر

```
skill: moussaif-mcp-bridge
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Adapter Coverage, Version Break Rate, Semantic Consistency
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 5 · الفصل 22 · التحقق والفشل والتكامل

M-AGIOS / 22

واجهات MCP وطبقة التوافق

MCP Interfaces and Compatibility Layer

VA · Validation, Failure Modes, and Integration

حالة تطبيق يمكن تبديل أداة Search بأخرى عبر MCP، لكن تظل قاعدة الكتاب أن الادعاء التقني يحتاج مصدرًا أوليًا واختبار حداثة مستقلة.

Protocol التجربة

1. ابن Baseline يعالج واجهات MCP وطبقة التوافق من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل اكتشاف قدرات وقت التشغيل.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- اعتبار Schema معنى كاملًا: اختبار مدخلات مضادة.
- صلاحيات Transport واسعة: اختبار مصدر أو أداة مستقلة.
- ترقية إصدار صامتة: اختبار استرداد وحدود.

مقاييس الفصل

Adapter Coverage	مقياس النتيجة الأساسية.
Version Break Rate	مقياس جودة المسار أو الدليل.
Semantic Consistency	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R26 · Model Context Protocol · MCP Specification 25-11-2025 | R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون اكتشاف قدرات وقت التشغيل
- من دون تطبيع الأخطاء والمهل
- من دون Moussaif MCP Bridge
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 5 · الفصل 23 · الأطروحة والحدود

M-AGIOS / 23

الوكيل المتصفح واستعمال الحاسوب

Browser and Computer-Use Agency

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل واجهات GUI جزئية الرؤية ومتغيرة؛ وقد يضغط الوكيل زرًا خاطئًا أو يكرر إرسال نموذج بسبب تأخر الصفحة.

الأطروحة المساعفية

يبني M-AGIOS Perception-Action Loop يفصل Observation عن Targeting، ويستخدم Anchors وحواجز قبل الأفعال الخارجية وتحققًا بعد كل خطوة حرجة.

الحد البنيوي الأدنى

يجب أن تظهر تثبيت الهدف بمعرفات متعددة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتجسيد المرئي

Moussaif Visual-Embodiment Theory

تقرر أن تثبيت الهدف بمعرفات متعددة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الوكيل المتصفح واستعمال الحاسوب. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للنقرة المفسّرة

Moussaif Explainable-Click Principle

تربط استخدام انتظار شرطي لا زمني فقط بـDuplicate Submission، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم جعل الفعل الخارجي عقدًا قابلاً للتحقق والعكس، لا تخمينًا لغويًا غير محكوم.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Visual-Embodiment Theory ليس دليلًا على وجود الآلية
- تنفيذ تثبيت الهدف بمعرفات متعددة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Action Grounding
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من استخدام انتظار شرطي لا زمني فقط الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 5 · الفصل 23 · الآلية والنموذج التشغيلي

M-AGIOS / 23

الوكيل المتصفح واستعمال الحاسوب

Browser and Computer-Use Agency

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تثبيت الهدف بمعرفات متعددة	استخدام انتظار شرطي لا زمني فقط	لقطة حالة قبل الإرسال والحذف

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يبني M-AGIOS Perception-Action Loop يفصل Observation عن Targeting، ويستخدم Anchors وحواجز قبل الأفعال الخارجية وتحققًا بعد كل خطوة حرجة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

A_gui = Ground(target, obs) × Guard(effect) × Verify(next_obs)

آلة الحالة المصغرة

Input	Goal Contract وواجهات GUI جزئية الرؤية ومتغيرة
Control	تثبيت الهدف بمعرفات متعددة
Transform	استخدام انتظار شرطي لا زمني فقط
Evidence	لقطة حالة قبل الإرسال والحذف
Stop	بلوغ عتبة Action Grounding أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تثبيت الهدف بمعرفات متعددة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Duplicate Submission فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 5 · الفصل 23 · الأدوات وعقود التنفيذ

M-AGIOS / 23

الوكيل المتصفح واستعمال الحاسوب

Browser and Computer-Use Agency

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُثبت الهدف المرئي

Visual Target Anchor

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تثبيت الهدف بمعرفات متعددة

المخرج: Artifact منظم يفعّل استخدام انتظار شرطي لا زمني فقط

الحارس: لا تتجاوز خطر انجراف Layout.

الأداة 2 · حارس الفعل الخارجي

External-Action Guard

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Recovery Steps

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Recovery Steps.

Runtime Contract المختصر

```
skill: visual-target-anchor
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Action Grounding, Duplicate Submission, Recovery Steps
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الوكيل المتصفح واستعمال الحاسوب
Browser and Computer-Use Agency

VA · Validation, Failure Modes, and Integration

حالة تطبيق قبل ضغط Send، يعيد النظام قراءة المستلم والموضوع وعدد المرفقات من الشاشة، ثم يتحقق من وجود الرسالة في Sent مرة واحدة.

Protocol التجربة

1. ابن Baseline يعالج الوكيل المتصفح واستعمال الحاسوب من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تثبيت الهدف بمعرفات متعددة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- انجراف Layout: اختبار مدخلات مضادة.
- Prompt Injection مرئي: اختبار مصدر أو أداة مستقلة.
- حالة تحميل ملتبسة: اختبار استرداد وحدود.

مقاييس الفصل

Action Grounding	مقياس النتيجة الأساسية.
Duplicate Submission	مقياس جودة المسار أو الدليل.
Recovery Steps	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R17 · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments · Tianbao Xie et al. · R20 | WebArena: A Realistic Web Environment for Building Autonomous Agents · Shuyan Zhou et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تثبيت الهدف بمعرفات متعددة
- من دون استخدام انتظار شرطي لا زمني فقط
- من دون Visual Target Anchor
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 5 · الفصل 24 · الأطروحة والحدود

M-AGIOS / 24

تنفيذ الكود والحوسبة الحتمية

Code Execution and Deterministic Compute

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُستخدم النموذج للحساب أو التحويل الذي تستطيع أداة حتمية إثباته، أو يُنفذ كود مولد مباشرة في بيئة واسعة.

الأطروحة المساعفية

يفصل Runtime بين Synthesis وExecution وVerification؛ يشغّل الكود في Sandbox بأقل صلاحيات، ويحتفظ بالمخرجات والاختبارات كدليل مستقل.

الحد البنيوي الأدنى

يجب أن تظهر تحديد I/O قبل كتابة الكود في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتحويل اللغة إلى برهان تنفيذي

Moussaif Executable-Proof Theory

تقرر أن تحديد I/O قبل كتابة الكود ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تنفيذ الكود والحوسبة الحتمية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لأفضل أداة حتمية

Moussaif Deterministic-Tool Preference

تربط Static Checks ثم Sandbox بـSandbox Escape Surface، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم جعل الفعل الخارجي عقدًا قابلاً للتحقق والعكس، لا تخمينًا لغويًا غير محكوم.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Executable-Proof Theory ليس دليلًا على وجود الآلية
- تنفيذ تحديد I/O قبل كتابة الكود يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Test Pass Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من Static Checks ثم Sandbox الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

تنفيذ الكود والحوسبة الحتمية

Code Execution and Deterministic Compute

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تحديد I/O قبل كتابة الكود	Sandbox ثم Static Checks	اختبار خصائص لا أمثلة فقط

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يفصل Runtime بين Synthesis وExecution وVerification؛ يشغل الكود في Sandbox بأقل صلاحيات، ويحتفظ بالمخرجات والاختبارات كدليل مستقل. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$\text{Trust(code)} = \text{Tests_independent} \times \text{Sandbox} \times \text{Reproducibility}$$

آلة الحالة المصغرة

Input	Goal Contract ويُستخدم النموذج للحساب أو التحويل الذي تستطيع أداة حتمية إثباته، أو يُنفذ كود مولد مباشرة في بيئة واسعة.
Control	تحديد I/O قبل كتابة الكود
Transform	Sandbox ثم Static Checks
Evidence	اختبار خصائص لا أمثلة فقط
Stop	بلوغ عتبة Test Pass Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تحديد I/O قبل كتابة الكود وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Sandbox Escape Surface فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 5 · الفصل 24 · الأدوات وعقود التنفيذ

M-AGIOS / 24

تنفيذ الكود والحوسبة الحتمية

Code Execution and Deterministic Compute

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · قفص التنفيذ

Execution Sandbox

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تحديد I/O قبل كتابة الكود

المخرج: Artifact منظم يفعّل Static Checks ثم Sandbox

الحارس: لا تتجاوز خطر اعتماد Tests كتبها المولد نفسه.

الأداة 2 · مولّد اختبار الخصائص

Property-Test Generator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Deterministic Coverage

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Deterministic Coverage.

Runtime Contract المختصر

```
skill: execution-sandbox
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Test Pass Rate, Sandbox Escape Surface, Deterministic Coverage
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

تنفيذ الكود والحوسبة الحتمية

Code Execution and Deterministic Compute

VA · Validation, Failure Modes, and Integration

حالة تطبيق يعد النظام صفحات PDF بوسيلة برمجية، ويقارن Metadata واستخراج النص بدل الثقة بأن المولد التزم بالعدد المطلوب.

Protocol التجربة

1. ابن Baseline يعالج تنفيذ الكود والحوسبة الحتمية من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تحديد I/O قبل كتابة الكود.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- اعتماد Tests كتبها المولد نفسه: اختبار مدخلات مضادة.
- وصول شبكة غير لازم: اختبار مصدر أو أداة مستقلة.
- أرقام صحيحة مصادفة: اختبار استرداد وحدود.

مقاييس الفصل

Test Pass Rate	مقياس النتيجة الأساسية.
Sandbox Escape Surface	مقياس جودة المسار أو الدليل.
Deterministic Coverage	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill، عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R10 · Timo Schick et al. · Toolformer: Language Models Can Teach Themselves to Use Tools | R21 · Carlos E. Jimenez et al. · SWE-bench: Can Language Models Resolve Real-World GitHub Issues?

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تحديد I/O قبل كتابة الكود
- من دون Static Checks ثم Sandbox
- من دون Execution Sandbox
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

M-AGIOS / 25

الإدراك متعدد الوسائط والمشغلات

Multimodal Perception and Actuators

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تصل الصورة والصوت والنص بأزمنة ودقة مختلفة، وقد تُدمج كأنها ملاحظات متزامنة ومتكافئة الثقة.

الأطروحة المساعفية

يعطي النظام كل Observation طابع زمن ومصدرًا ودقة ونطاقًا، ثم يبني Fusion Graph يصرح بالتعارض ويؤخر الفعل إذا لم يبلغ الدليل عتبته.

الحد البنيوي الأدنى

يجب أن تظهر معايرة كل Sensor منفردًا في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للاندماج الوعائي

Moussaif Vascular Sensor Fusion

تقرر أن معايرة كل Sensor منفردًا ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الإدراك متعدد الوسائط والمشغلات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لعتبة التجسيد

Moussaif Embodiment-Threshold Law

تربط محاذاة زمنية قبل الدمج بـTemporal Alignment، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم جعل الفعل الخارجي عقدًا قابلاً للتحقق والعكس، لا تخمينًا لغويًا غير محكوم.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Vascular Sensor Fusion ليس دليلًا على وجود الآلية
- تنفيذ معايرة كل Sensor منفردًا يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Fusion Calibration
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من محاذاة زمنية قبل الدمج الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

الإدراك متعدد الوسائط والمشغلات

Multimodal Perception and Actuators

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
معايرة كل Sensor منفردًا	محاذاة زمنية قبل الدمج	اختبار Cross-Modal Consistency

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يعطي النظام كل Observation طابع زمن ومصدرا ودقة ونطاقا، ثم يبني Fusion Graph يصرح بالتعارض ويؤخر الفعل إذا لم يبلغ الدليل عتبته. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$F_{obs} = \Sigma w_i O_i, \text{ where } w_i = reliability_i \times freshness_i$$

آلة الحالة المصغرة

Input	Goal Contract وتصل الصورة والصوت والنص بأزمنة ودقة مختلفة
Control	معايرة كل Sensor منفردًا
Transform	محاذاة زمنية قبل الدمج
Evidence	اختبار Cross-Modal Consistency
Stop	بلوغ عتبة Fusion Calibration أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت معايرة كل Sensor منفردًا وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Temporal Alignment فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 5 · الفصل 25 · الأدوات وعقود التنفيذ

M-AGIOS / 25

الإدراك متعدد الوسائط والمشغلات

Multimodal Perception and Actuators

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · موازن ثقة الحواس

Sensor-Confidence Balancer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، معايرة كل Sensor منفردًا

المخرج: Artifact منظم يفعّل محاذاة زمنية قبل الدمج

الحارس: لا تتجاوز خطر هلوسة وصف الصورة.

الأداة 2 · كاشف التعارض متعدد الوسائط

Cross-Modal Conflict Detector

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Unsafe Actuation Rate

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Unsafe Actuation Rate.

Runtime Contract المختصر

```
skill: sensor-confidence-balancer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Fusion Calibration, Temporal Alignment, Unsafe Actuation Rate
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الإدراك متعدد الوسائط والمشغلات

Multimodal Perception and Actuators

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا قرأ OCR رقمًا يختلف عن النص البنيوي، لا يرسل النظام الدفعة؛ يطلب قناة ثالثة أو مراجعة بشرية بحسب حجم الأثر.

Protocol التجربة

1. ابن Baseline يعالج الإدراك متعدد الوسائط والمشغلات من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل معايرة كل Sensor منفردًا.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- هلوسة وصف الصورة: اختبار مدخلات مضادة.
- صوت متأخر: اختبار مصدر أو أداة مستقلة.
- فعل سريع قبل اكتمال الإدراك: اختبار استرداد وحدود.

مقاييس الفصل

Fusion Calibration	مقياس النتيجة الأساسية.
Temporal Alignment	مقياس جودة المسار أو الدليل.
Unsafe Actuation Rate	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R20 · Tianbao Xie et al. · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments | R25 · NIST · Artificial Intelligence Risk Management Framework (1.0 AI RMF)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون معايرة كل Sensor منفردًا
- من دون محاذاة زمنية قبل الدمج
- من دون Sensor-Confidence Balancer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 6 · الفصل 26 · الأطروحة والحدود

M-AGIOS / 26

تشريح HyperSkill المساعفية

Anatomy of a Moussaif HyperSkill

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل ملفات المهارات غالبًا تعليمات عامة لا تحدد شروط التشغيل أو أدلة النجاح، فتحتاج إعادة تفسير كاملة في كل مرة.

الأطروحة المساعفية

HyperSkill وحدة موقّعة تضم Purpose وTriggers وInputs وPreconditions وProcedure وTools وVerification وFailure Modes وCost Profile وVersion.

الحد البنيوي الأدنى

يجب أن تظهر تعريف Interface ثابت في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتعبئة القدرة

Moussaif Capability-Packaging Theory

تقرر أن تعريف Interface ثابت ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تشريح HyperSkill المساعفية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف لاكتمال المهارة

Moussaif Skill-Completeness Principle

تربط فصل الإجراء عن أمثلة المجال بـContract Completeness، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعبئة الخبرة في وحدات تشغيلية قابلة للقياس وإعادة الاستخدام والتطوير الآمن.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Capability-Packaging Theory ليس دليلًا على وجود الآلية
- تنفيذ تعريف Interface ثابت يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Reuse Success
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من فصل الإجراء عن أمثلة المجال الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 6 · الفصل 26 · الآلية والنموذج التشغيلي

M-AGIOS / 26

تشريح HyperSkill المساعفية

Anatomy of a Moussaif HyperSkill

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تعريف Interface ثابت	فصل الإجراء عن أمثلة المجال	إرفاق اختبارات قابلة للتنفيذ

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. HyperSkill وحدة مُوقّعة تضم Purpose وTriggers وInputs وPreconditions وProcedure وTools وVerification وFailure Modes وCost Profile وVersion. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

SkillReady = Interface × Preconditions × Procedure × Verification × Recovery

آلة الحالة المصغرة

Input	Goal Contract وملفات المهارات غالبًا تعليمات عامة لا تحدد شروط التشغيل أو أدلة النجاح، فتحتاج إعادة تفسير كاملة في كل مرة.
Control	تعريف Interface ثابت
Transform	فصل الإجراء عن أمثلة المجال
Evidence	إرفاق اختبارات قابلة للتنفيذ
Stop	بلوغ عتبة Reuse Success أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تعريف Interface ثابت وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Contract Completeness فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 6 · الفصل 26 · الأدوات وعقود التنفيذ

M-AGIOS / 26

تشريح HyperSkill المساعفية

Anatomy of a Moussaif HyperSkill

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُنشئ HyperSkill

HyperSkill Builder

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تعريف Interface ثابت

المخرج: Artifact منظم يفعّل فصل الإجراء عن أمثلة المجال

الحارس: لا تتجاوز خطر تعليمات بلا Preconditions.

الأداة 2 · محص اكتمال المهارة

Skill-Completeness Auditor

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Invocation Cost

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Invocation Cost.

Runtime Contract المختصر

```
skill: hyperskill-builder
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Reuse Success, Contract Completeness, Invocation Cost
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 6 · الفصل 26 · التحقق والفشل والتكامل

M-AGIOS / 26

تشريح HyperSkill المساعفية

Anatomy of a Moussaif HyperSkill

VA · Validation, Failure Modes, and Integration

حالة تطبيق مهارة ترحيل نطاق لا تقول فقط استبدل الرابط؛ بل تجرد المستودعات، تستثني الأرشيف، تفحص الروابط المشفرة، تشغل الاختبارات وتثبت غياب القديم.

Protocol التجربة

1. ابنِ Baseline يعالج تشريح HyperSkill المساعفية من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تعريف Interface ثابت.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تعليمات بلا Preconditions: اختبار مدخلات مضادة.
- مثال يختلط بالقاعدة: اختبار مصدر أو أداة مستقلة.
- Version بلا اختبارات: اختبار استرداد وحدود.

مقاييس الفصل

Reuse Success	مقياس النتيجة الأساسية.
Contract Completeness	مقياس جودة المسار أو الدليل.
Invocation Cost	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R16 · Qingyun Wu et al. · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation | R21 · Carlos E. Jimenez et al. · SWE-bench: Can Language Models Resolve Real-World GitHub Issues?

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تعريف Interface ثابت
- من دون فصل الإجراء عن أمثلة المجال
- من دون HyperSkill Builder
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 6 · الفصل 27 · الأطروحة والحدود

M-AGIOS / 27

الحلقة الفائقة المحدودة

Bounded Hyperloop

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تُطرح Self-Improvement كحلقة مفتوحة؛ فتكرر التفكير أو تغيّر هدفها أو تستهلك الميزانية من دون معيار زيادة حقيقية.

الأطروحة المساعفية

Bounded Hyperloop تكرر دورة Hypothesis وAction وObservation وFalsification وUpdate تحت حدود Iteration وCost وRisk وEvidence Gain.

الحد البنيوي الأدنى

يجب أن تظهر ميزانية تكرار صريحة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للحلقة المحدودة

Moussaif Bounded-Hyperloop Theory

تقرر أن ميزانية تكرار صريحة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الحلقة الفائقة المحدودة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لحفظ أفضل حالة

Moussaif Best-State Preservation

تربط تسجيل أفضل حالة معروفة بـLoop Cost، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعبئة الخبرة في وحدات تشغيلية قابلة للقياس وإعادة الاستخدام والتطوير الآمن.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Bounded-Hyperloop Theory ليس دليلًا على وجود الآلية
- تنفيذ ميزانية تكرار صريحة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Evidence Gain per Iteration
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تسجيل أفضل حالة معروفة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف نتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 6 · الفصل 27 · الآلية والنموذج التشغيلي

M-AGIOS / 27

الحلقة الفائقة المحدودة

Bounded Hyperloop

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
ميزانية تكرار صريحة	تسجيل أفضل حالة معروفة	توقف عند تشبع الدليل أو طلب إذن

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. Bounded Hyperloop تكرر دورة Hypothesis وAction وObservation وFalsification وUpdate تحت حدود Iteration وCost وRisk وEvidence Gain. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Loop_{i+1} only if ΔEvidence_i > Cost_i + Risk_i

آلة الحالة المصغرة

Input	Goal Contract وتُطرح Self-Improvement كحلقة مفتوحة
Control	ميزانية تكرار صريحة
Transform	تسجيل أفضل حالة معروفة
Evidence	توقف عند تشبع الدليل أو طلب إذن
Stop	بلوغ عتبة Evidence Gain per Iteration أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت ميزانية تكرار صريحة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Loop Cost فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 6 · الفصل 27 · الأدوات وعقود التنفيذ

M-AGIOS / 27

الحلقة الفائقة المحدودة

Bounded Hyperloop

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · متحكم Hyperloop

Hyperloop Controller

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، ميزانية تكرار صريحة

المخرج: Artifact منظم يفعّل تسجيل أفضل حالة معروفة

الحارس: لا تتجاوز خطر تأمل بلا اختبار.

الأداة 2 · مسجل أفضل حالة

Best-State Register

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Regression Avoidance

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Regression Avoidance.

Runtime Contract المختصر

```
skill: hyperloop-controller
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Evidence Gain per Iteration, Loop Cost, Regression Avoidance
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 6 · الفصل 27 · التحقق والفشل والتكامل

M-AGIOS / 27

الحلقة الفائقة المحدودة

Bounded Hyperloop

VA · Validation, Failure Modes, and Integration

حالة تطبيق في تحسين Prompt، يحتفظ النظام بنسخة Benchmark الأفضل، ولا يقبل إصدارًا أحدث لمجرد أنه أطول أو أكثر تعقيدًا.

Protocol التجربة

1. ابنِ Baseline يعالج الحلقة الفائقة المحدودة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل ميزانية تكرار صريحة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تأمل بلا اختبار: اختبار مدخلات مضادة.
- نسيان أفضل نسخة: اختبار مصدر أو أداة مستقلة.
- زيادة الاستقلال أثناء الفشل: اختبار استرداد وحدود.

مقاييس الفصل

Evidence Gain per Iteration	مقياس النتيجة الأساسية.
Loop Cost	مقياس جودة المسار أو الدليل.
Regression Avoidance	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R11 · Noah Shinn et al. · Reflexion: Language Agents with Verbal Reinforcement Learning | R12 · Shunyu Yao et al. · Tree of Thoughts: Deliberate Problem Solving with Large Language Models

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون ميزانية تكرار صريحة
- من دون تسجيل أفضل حالة معروفة
- من دون Hyperloop Controller
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة.
Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 6 · الفصل 28 · الأطروحة والحدود

M-AGIOS / 28

تولید المهارات من الفجوات

Skill Synthesis from Capability Gaps

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل قد ينتج النظام مهارة جديدة لكل مهمة، فيتضاعف المستودع بتكرارات طفيفة ولا تُعرف الحاجة الحقيقية.

الأطروحة المساعفية

لا تُولد Skill إلا بعد Gap Analysis يثبت غياب تركيب قائم اقتصادي، ثم تُبنى من Episodes متعددة وتدخل Sandbox واختبارات حدودية.

الحد البنيوي الأدنى

يجب أن تظهر البحث عن تركيب قبل الإنشاء في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتبلور المهارة

Moussaif Skill-Crystallization Theory

تقرر أن البحث عن تركيب قبل الإنشاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توليد المهارات من الفجوات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لحداثة القدرة

Moussaif Capability-Novelty Law

تربط تجميع أمثلة نجاح وفشل بـDuplicate Skill Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعبئة الخبرة في وحدات تشغيلية قابلة للقياس وإعادة الاستخدام والتطوير الآمن.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Skill-Crystallization Theory ليس دليلًا على وجود الآلية
- تنفيذ البحث عن تركيب قبل الإنشاء يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Novel Skill Yield
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تجميع أمثلة نجاح وفشل الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

M-AGIOS / 28

توليد المهارات من الفجوات

Skill Synthesis from Capability Gaps

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
البحث عن تركيب قبل الإنشاء	تجميع أمثلة نجاح وفشل	إصدار Candidate غير مفعل حتى التحقق

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. لا تُولد Skill إلا بعد Gap Analysis يثبت غياب تركيب قائم اقتصادي، ثم تبنى من Episodes متعددة وتدخل Sandbox واختبارات حدودية. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

CreateSkill iff GapVerified × ReusePotential × Testability > Threshold

آلة الحالة المصغرة

Input	Goal Contract وقد ينتج النظام مهارة جديدة لكل مهمة، فيتضاعف المستودع بتكرارات طفيفة ولا تعرف الحاجة الحقيقية.
Control	البحث عن تركيب قبل الإنشاء
Transform	تجميع أمثلة نجاح وفشل
Evidence	إصدار Candidate غير مفعل حتى التحقق
Stop	بلوغ عتبة Novel Skill Yield أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت البحث عن تركيب قبل الإنشاء وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Duplicate Skill Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 6 · الفصل 28 · الأدوات وعقود التنفيذ

M-AGIOS / 28

توليد المهارات من الفجوات

Skill Synthesis from Capability Gaps

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · محكمة فجوة القدرة

Capability-Gap Court

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، البحث عن تركيب قبل الإنشاء

المخرج: Artifact منظم يفعّل تجميع أمثلة نجاح وفشل

الحارس: لا تتجاوز خطر تسمية جديدة لسلوك قديم.

الأداة 2 · مُبلور المهارة

Skill Crystallizer

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Boundary-Test Pass

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Boundary-Test Pass.

Runtime Contract المختصر

```
skill: capability-gap-court
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Novel Skill Yield, Duplicate Skill Rate, Boundary-Test Pass
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 6 · الفصل 28 · التحقق والفشل والتكامل

M-AGIOS / 28

توليد المهارات من الفجوات

Skill Synthesis from Capability Gaps

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا تكرر فشل ضبط اتجاه RTL في ثلاثة مولدات، يبني النظام مهارة Typst عربية بمتطلبات Font وDirection وRender QA بدل Prompt عام للوثائق.

Protocol التجربة

1. ابنِ Baseline يعالج توليد المهارات من الفجوات من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل البحث عن تركيب قبل الإنشاء.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تسمية جديدة لسلوك قديم: اختبار مدخلات مضادة.
- تعلم من Episode واحد: اختبار مصدر أو أداة مستقلة.
- نشر Skill غير موقعة: اختبار استرداد وحدود.

مقاييس الفصل

Novel Skill Yield	مقياس النتيجة الأساسية.
Duplicate Skill Rate	مقياس جودة المسار أو الدليل.
Boundary-Test Pass	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R11 · Noah Shinn et al. · Reflexion: Language Agents with Verbal Reinforcement Learning | R27 · Wujiang Xu et al. · A-Mem: Agentic Memory for LLM Agents

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون البحث عن تركيب قبل الإنشاء
- من دون تجميع أمثلة نجاح وفشل
- من دون Capability-Gap Court
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 6 · الفصل 29 · الأطروحة والحدود

M-AGIOS / 29

تقييم المهارات وإصدارها

Skill Evaluation and Versioning

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تتغير مهارة لأن نصها تحسن أسلوبيًا، لكن لا يُقاس أثرها ولا يُعرف إن كسرت حالات قديمة.

الأطروحة المساعفية

يعامل Skill Version كإصدار برمجي: Semantic Versioning وظيفي، Benchmarks ثابتة، Canary، سجل أسباب، وRollback إلى آخر نسخة موثوقة.

الحد البنيوي الأدنى

يجب أن تظهر فصل تغيير العقد عن تحسين داخلي في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لنضج المهارة

Moussaif Skill-Maturity Theory

تقرر أن فصل تغيير العقد عن تحسين داخلي ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تقييم المهارات وإصدارها. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للإصدار القابل للرجوع

Moussaif Revertible-Version Principle

تربط اختبار Regression ومجالات جديدة بـVersion Adoption، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعبئة الخبرة في وحدات تشغيلية قابلة للقياس وإعادة الاستخدام والتطوير الآمن.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Skill-Maturity Theory ليس دليلًا على وجود الآلية
- تنفيذ فصل تغيير العقد عن تحسين داخلي يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Regression Delta
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من اختبار Regression ومجالات جديدة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 6 · الفصل 29 · الآلية والنموذج التشغيلي

M-AGIOS / 29

تقييم المهارات وإصدارها

Skill Evaluation and Versioning

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل تغيير العقد عن تحسين داخلي	اختبار Regression ومجالات جديدة	تقدير الثقة بالسياق لا المتوسط فقط

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يعامل Skill Version كإصدار برمجي: Semantic Versioning وظيفي، Benchmarks ثابتة، Canary، سجل أسباب، وRollback إلى آخر نسخة موثوقة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Promote(v) iff ΔQuality - ΔCost - ΔRisk > 0 across required strata

آلة الحالة المصغرة

Input	Goal Contract وتغير مهارة لأن نصها تحسن أسلوبيًا، لكن لا يقاس أثرها ولا يعرف إن كسرت حالات قديمة.
Control	فصل تغيير العقد عن تحسين داخلي
Transform	اختبار Regression ومجالات جديدة
Evidence	تقدير الثقة بالسياق لا المتوسط فقط
Stop	بلوغ عتبة Regression Delta أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل تغيير العقد عن تحسين داخلي وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Version Adoption فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 6 · الفصل 29 · الأدوات وعقود التنفيذ

M-AGIOS / 29

تقييم المهارات وإصدارها

Skill Evaluation and Versioning

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مقعد اختبار HyperSkill

HyperSkill Test Bench

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل تغيير العقد عن تحسين داخلي

المخرج: Artifact منظم يفعّل اختبار Regression ومجالات جديدة

الحارس: لا تتجاوز خطر Benchmark ملوث.

الأداة 2 · مسجل نسب المهارة

Skill-Lineage Registry

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Contextual Reliability

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Contextual Reliability.

Runtime Contract المختصر

```
skill: hyperskill-test-bench
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Regression Delta, Version Adoption, Contextual Reliability
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

تقييم المهارات وإصدارها
Skill Evaluation and Versioning

VA · Validation, Failure Modes, and Integration

حالة تطبيق يُطرح الإصدار الجديد على 10% من المهام منخفضة الأثر؛ إن انخفضت دقة التحقق يُعاد القديم تلقائيًا مع حفظ الأدلة.

Protocol التجربة

1. ابنِ Baseline يعالج تقييم المهارات وإصدارها من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل فصل تغيير العقد عن تحسين داخلي.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Benchmark ملوث: اختبار مدخلات مضادة.
- ترقية شاملة مباشرة: اختبار مصدر أو أداة مستقلة.
- إخفاء حالات الفشل النادرة: اختبار استرداد وحدود.

مقاييس الفصل

Regression Delta	مقياس النتيجة الأساسية.
Version Adoption	مقياس جودة المسار أو الدليل.
Contextual Reliability	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R18 · Xiao Liu et al. · AgentBench: Evaluating LLMs as Agents · R21 | Carlos E. Jimenez et al. · SWE-bench: Can Language Models Resolve Real-World GitHub Issues?

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل تغيير العقد عن تحسين داخلي
- من دون اختبار Regression ومجالات جديدة
- من دون HyperSkill Test Bench
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 6 · الفصل 30 · الأطروحة والحدود

M-AGIOS / 30

الانتشار التناسلي المحكوم

Governed Propagative Deployment

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل نسخ المهارات وKB إلى كل مشروع يخلق تباعدًا وإصدارات مجهولة، وقد تنشر قاعدة خاصة في سياق غير مصرح.

الأطروحة المساعفية

Moussaif Propagative Deployment ينشر References موقعة وحزمًا متخصصة وفق Project Map، مع سياسات Scope وCompatibility وRevocation.

الحد البنيوي الأدنى

يجب أن تظهر سجل أصل واحد Lineage في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للانتشار التناسلي

Moussaif Propagative-Deployment Theory

تقرر أن سجل أصل واحد Lineage ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الانتشار التناسلي المحكوم. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للوراثة المقيدة

Moussaif Constrained-Inheritance Law

تربط حزمة مشروع بأقل قدر بـUnauthorized Propagation، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعبئة الخبرة في وحدات تشغيلية قابلة للقياس وإعادة الاستخدام والتطوير الآمن.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Propagative-Deployment Theory ليس دليلًا على وجود الآلية
- تنفيذ سجل أصل واحد Lineage يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Version Convergence
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من حزمة مشروع بأقل قدر الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

الانتشار التناسلي المحكوم
Governed Propagative Deployment

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
سجل أصل واحد Lineage	حزمة مشروع بأقل قدر	قناة سحب التحديث والإبطال

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. Moussaif Propagative Deployment ينشر References موقعة وحزماً متخصصة وفق Project Map، مع سياسات Scope وCompatibility وRevocation. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

DeployPack(p) = CommonKernel + AuthorizedDomain + CompatibleSkills - PrivateState

آلة الحالة المصغرة

Input	Goal Contract ونسخ المهارات وKB إلى كل مشروع يخلق تباعداً وإصدارات مجهولة، وقد تنشر قاعدة خاصة في سياق غير مصرح.
Control	سجل أصل واحد Lineage
Transform	حزمة مشروع بأقل قدر
Evidence	قناة سحب التحديث والإبطال
Stop	بلوغ عتبة Convergence Version أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت سجل أصل واحد Lineage وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. واذا تحسن Outcome وانهار Unauthorized Propagation فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 6 · الفصل 30 · الأدوات وعقود التنفيذ

M-AGIOS / 30

الانتشار التناسلي المحكوم

Governed Propagative Deployment

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · موجّه الانتشار

Propagation Router

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، سجل أصل واحد Lineage

المخرج: Artifact منظم يفعّل حزمة مشروع بأقل قدر

الحارس: لا تتجاوز خطر نسخ كامل.

الأداة 2 · مراقب انحراف الحزم

Pack-Drift Monitor

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Revocation Latency

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Revocation Latency.

Runtime Contract المختصر

```
skill: propagation-router
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Version Convergence, Unauthorized Propagation, Revocation Latency
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الانتشار التناسلي المحكوم
Governed Propagative Deployment

VA · Validation, Failure Modes, and Integration

حالة تطبيق يتلقى مشروع جديد دستورًا عامًا وKB المجال وثلاث مهارات متوافقة فقط؛ لا يرث أسرار مزود أو سجل عملاء من مشروع آخر.

Protocol التجربة

1. ابنِ Baseline يعالج الانتشار التناسلي المحكوم من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل سجل أصل واحد Lineage.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- نسخ كامل: اختبار مدخلات مضادة.
- تسرب سياق مشروع: اختبار مصدر أو أداة مستقلة.
- تحديث يكسر Adapter: اختبار استرداد وحدود.

مقاييس الفصل

Version Convergence	مقياس النتيجة الأساسية.
Unauthorized Propagation	مقياس جودة المسار أو الدليل.
Revocation Latency	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R16 · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation · Qingyun Wu et al. · MCP · Model Context Protocol · 25-11-2025 Specification

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون سجل أصل واحد Lineage
- من دون حزمة مشروع بأقل قدر
- من دون Propagation Router
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 7 · الفصل 31 · الأطروحة والحدود

M-AGIOS / 31

متى نستخدم الوكلاء الفرعيين

When to Use Subagents

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُفترض أن تعدد الوكلاء أفضل دائمًا، مع أن الاتصال وازدواج العمل وتعارض السياق قد تجعل وكيلًا واحدًا أكثر دقة وأقل كلفة.

الأطروحة المساعفية

يفوض M-AGIOS فقط Subtask مستقلة ذات Interface ودليل إكمال عندما تتجاوز قيمة التوازي أو التخصص كلفة التنسيق والمخاطرة.

الحد البنيوي الأدنى

يجب أن تظهر تحليل Dependency Graph في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لعتبة التفويض

Moussaif Delegation-Threshold Theory

تقرر أن تحليل Dependency Graph ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح متى نستخدم الوكلاء الفرعيين. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للعدد الأدنى الكافي

Moussaif Minimum-Sufficient-Agent Law

تربط تقدير Coordination Tax بـ Coordination Tax، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم توزيع العمل فقط عندما تتجاوز مكاسب التخصص كلفة الاتصال والتضارب.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Delegation-Threshold Theory ليس دليلًا على وجود الآلية
- تنفيذ تحليل Dependency Graph يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Parallel Speedup
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تقدير Coordination Tax الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 7 · الفصل 31 · الآلية والنموذج التشغيلي

M-AGIOS / 31

متى نستخدم الوكلاء الفرعيين

When to Use Subagents

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تحليل Dependency Graph	تقدير Coordination Tax	اختيار أقل عدد من الوكلاء

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يفوض M-AGIOS فقط Subtask مستقلة ذات Interface ودليل إكمال عندما تتجاوز قيمة التوازي أو التخصص كلفة التنسيق والمخاطرة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Delegate iff Gain_parallel + Gain_specialty > Cost_coord + Risk_merge

آلة الحالة المصغرة

Input	Goal Contract ويُفترض أن تعدد الوكلاء أفضل دائمًا، مع أن الاتصال وازدواج العمل وتعارض السياق قد تجعل وكيلًا واحدًا أكثر دقة وأقل كلفة.
Control	تحليل Dependency Graph
Transform	تقدير Coordination Tax
Evidence	اختيار أقل عدد من الوكلاء
Stop	بلوغ عتبة Parallel Speedup أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تحليل Dependency Graph وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Coordination Tax فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 7 · الفصل 31 · الأدوات وعقود التنفيذ

M-AGIOS / 31

متى نستخدم الوكلاء الفرعيين

When to Use Subagents

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · حاسبة قيمة التفويض

Delegation-Value Calculator

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تحليل Dependency Graph

المخرج: Artifact منظم يفعّل تقدير Coordination Tax

الحارس: لا تتجاوز خطر تفويض مهمة مترابطة.

الأداة 2 · قاطع التكرار الوكيلي

Agent-Duplication Cutter

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Duplicate Work

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Duplicate Work.

Runtime Contract المختصر

```
skill: delegation-value-calculator
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Parallel Speedup, Coordination Tax, Duplicate Work
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

متى نستخدم الوكلاء الفرعيين
When to Use Subagents

VA · Validation, Failure Modes, and Integration

حالة تطبيق يمكن تفويض مراجعة المصادر وتصميم Benchmark بالتوازي، لكن دمج النظرية النهائية يبقى عند Orchestrator واحد يحفظ الاتساق الاصطلاحي.

Protocol التجربة

1. ابنِ Baseline يعالج متى نستخدم الوكلاء الفرعيين من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تحليل Dependency Graph.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تفويض مهمة مترابطة: اختبار مدخلات مضادة.
- سياق ناقص: اختبار مصدر أو أداة مستقلة.
- إطلاق وكلاء أكثر من موارد التحقق: اختبار استرداد وحدود.

مقاييس الفصل

Parallel Speedup	مقياس النتيجة الأساسية.
Coordination Tax	مقياس جودة المسار أو الدليل.
Duplicate Work	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R16 · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation · Qingyun Wu et al. · R30 | ChatDev: Communicative Agents for Software Development · Chen Qian et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون Dependency Graph تحليل
- من دون Coordination Tax تقدير
- من دون Delegation-Value Calculator
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 7 · الفصل 32 · الأطروحة والحدود

M-AGIOS / 32

توليد الأدوار والعقود بين الوكلاء

Role Synthesis and Inter-Agent Contracts

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تُعطى الوكلاء أسماء مثل ناقد وخبير بلا حدود مخرجات أو صلاحيات، فتتحول المحادثة إلى Role-play غير قابل للقياس.

الأطروحة المساعفية

يُشتق الدور من فجوة عمل محددة، ويحمل Input Contract وOutput Schema وEvidence Duty وDeadline وForbidden Actions وعلاقة تبعية.

الحد البنيوي الأدنى

يجب أن تظهر توليد الدور من Task Graph في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للدور المشتق

Moussaif Derived-Role Theory

تقرر أن توليد الدور من Task Graph ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توليد الأدوار والعقود بين الوكلاء. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للعقد البيني

Moussaif Inter-Agent Contract

تربط عزل Context حسب الحاجة بـEvidence Duty Fulfillment، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم توزيع العمل فقط عندما تتجاوز مكاسب التخصص كلفة الاتصال والتضارب.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Derived-Role Theory ليس دليلًا على وجود الآلية
- تنفيذ توليد الدور من Task Graph يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Output Schema Compliance
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من عزل Context حسب الحاجة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

توليد الأدوار والعقود بين الوكلاء

Role Synthesis and Inter-Agent Contracts

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
توليد الدور من **Task Graph**	عزل **Context** حسب الحاجة	إرجاع **Artifact** لا رأي عام

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يُشتق الدور من فجوة عمل محددة، ويحمل Input Contract وOutput Schema وEvidence Duty وDeadline وForbidden Actions وعلاقة تبعية. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Role = Gap + Inputs + Deliverable + Evidence + Limits + Deadline

آلة الحالة المصغرة

Input	Goal Contract وتُعطى الوكلاء أسماء مثل ناقد وخبير بلا حدود مخرجات أو صلاحيات، فتتحول المحادثة إلى Role-play غير قابل للقياس.
Control	توليد الدور من Task Graph
Transform	عزل Context حسب الحاجة
Evidence	إرجاع Artifact لا رأي عام
Stop	بلوغ عتبة Output Schema Compliance أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت توليد الدور من Task Graph وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Evidence Duty Fulfillment فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 7 · الفصل 32 · الأدوات وعقود التنفيذ

M-AGIOS / 32

توليد الأدوار والعقود بين الوكلاء

Role Synthesis and Inter-Agent Contracts

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُركب أدوار الوكلاء

Agent-Role Composer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، توليد الدور من Task Graph

المخرج: Artifact منظم يفعّل عزل Context حسب الحاجة

الحارس: لا تتجاوز خطر شخصنة الدور.

الأداة 2 · مدقق عقد التفويض

Delegation-Contract Auditor

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Role Overlap

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Role Overlap.

Runtime Contract المختصر

```
skill: agent-role-composer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Output Schema Compliance, Evidence Duty Fulfillment, Role Overlap
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

توليد الأدوار والعقود بين الوكلاء
Role Synthesis and Inter-Agent Contracts

VA · Validation, Failure Modes, and Integration

حالة تطبيق وكيل المصادر يعيد جدول Claim-Source-Date-Confidence لا فقرة رأي، ولا يمتلك صلاحية تعديل الكتاب أو اختيار النتيجة النظرية.

Protocol التجربة

1. ابن Baseline يعالج توليد الأدوار والعقود بين الوكلاء من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل توليد الدور من Task Graph.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- شخصنة الدور: اختبار مدخلات مضادة.
- صلاحية زائدة: اختبار مصدر أو أداة مستقلة.
- مخرج سردي غير قابل للدمج: اختبار استرداد وحدود.

مقاييس الفصل

Output Schema Compliance	مقياس النتيجة الأساسية.
Evidence Duty Fulfillment	مقياس جودة المسار أو الدليل.
Role Overlap	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R16 · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation · Qingyun Wu et al. · R30 | ChatDev: Communicative Agents for Software Development · Chen Qian et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون توليد الدور من Task Graph
- من دون عزل Context حسب الحاجة
- من دون Agent-Role Composer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 7 · الفصل 33 · الأطروحة والحدود

M-AGIOS / 33

التنسيق والتزامن ودمج الحالة

Coordination, Synchronization, and State Merge

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تصل نتائج الوكلاء في أزمنة مختلفة وقد تعتمد على نسخ حالة متباينة، فيدمج المنسق توصيات متعارضة أو قديمة.

الأطروحة المساعفية

يستخدم النظام Versioned State و:Merge Policies جمع للحقائق المستقلة، مقارنة للبدائل، رفض للكتابة المتعارضة، وإعادة تشغيل عند Stale Read.

الحد البنيوي الأدنى

يجب أن تظهر تثبيت Snapshot لكل تفويض في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للاتساق الوكيلي

Moussaif Agent-Consistency Theory

تقرر أن تثبيت Snapshot لكل تفويض ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التنسيق والتزامن ودمج الحالة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للقراءة القديمة

Moussaif Stale-Read Law

تربط تصنيف نوع الدمج بـStale Result Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم توزيع العمل فقط عندما تتجاوز مكاسب التخصص كلفة الاتصال والتضارب.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Agent-Consistency Theory ليس دليلًا على وجود الآلية
- تنفيذ تثبيت Snapshot لكل تفويض يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Merge Conflict Rate
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تصنيف نوع الدمج الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

التنسيق والتزامن ودمج الحالة

Coordination, Synchronization, and State Merge

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تثبيت Snapshot لكل تفويض	تصنيف نوع الدمج	كشف التعارض قبل تحديث World State

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يستخدم النظام Versioned State وMerge Policies: جمع للحقائق المستقلة، مقارنة للبدائل، رفض للكتابة المتعارضة، وإعادة تشغيل عند Stale Read. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

Merge(r_i) valid iff BaseVersion_i compatible with CurrentState

آلة الحالة المصغرة

Input	Goal Contract وتصل نتائج الوكلاء في أزمنة مختلفة وقد تعتمد على نسخ حالة متباينة، فيدمج المنسق توصيات متعارضة أو قديمة.
Control	تثبيت Snapshot لكل تفويض
Transform	تصنيف نوع الدمج
Evidence	كشف التعارض قبل تحديث World State
Stop	بلوغ عتبة Merge Conflict Rate أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تثبيت Snapshot لكل تفويض وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Stale Result Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

التنسيق والتزامن ودمج الحالة
Coordination, Synchronization, and State Merge

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُدمج الحالة الموقّتة
Versioned-State Merger

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.
المدخل: Goal Contract، الحالة الحالية، تثبيت Snapshot لكل تفويض
المخرج: Artifact منظم يفعّل تصنيف نوع الدمج
الحارس: لا تتجاوز خطر Last Writer Wins.

الأداة 2 · كاشف التعارض السببي
Causal-Conflict Detector

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة State Convergence
المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ
المقياس: State Convergence.

Runtime Contract المختصر

```
skill: versioned-state-merger
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Merge Conflict Rate, Stale Result Rate, State Convergence
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل
- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل
- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events
tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ
خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

التنسيق والتزامن ودمج الحالة
Coordination, Synchronization, and State Merge

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا بُني فصل على تعريف سابق ثم غيّره وكيل آخر، يُوسم الفصل Stale ويعاد التحقق بدل دمجه لأن صياغته جيدة.

Protocol التجربة

1. ابن Baseline يعالج التنسيق والتزامن ودمج الحالة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تثبيت Snapshot لكل تفويض.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Last Writer Wins: اختبار مدخلات مضادة.
- Provenance: فقد اختبار مصدر أو أداة مستقلة.
- اعتماد دائري بين وكلاء: اختبار استرداد وحدود.

مقاييس الفصل

Merge Conflict Rate	مقياس النتيجة الأساسية.
Stale Result Rate	مقياس جودة المسار أو الدليل.
State Convergence	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R16 · ChatDev: Communicative Agents for Software Development · Chen Qian et al. · R30 | AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation · Qingyun Wu et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تثبيت Snapshot لكل تفويض
- من دون تصنيف نوع الدمج
- من دون Versioned-State Merger
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 7 · الفصل 34 · الأطروحة والحدود

M-AGIOS / 34

المداولة والمحكمة متعددة الوكلاء

Multi-Agent Deliberation and Court

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل الاتفاق بين وكلاء متشابهين يوهم بالثقة، والجدل المفتوح يضاعف Tokens من دون استقلال أدلة أو قرار تحكيم.

الأطروحة المساعفية

محكمة مساعف توزع أدوار Claimant وFalsifier وEvidence Judge وRisk Judge، وتمنع كل دور من تغيير معيار الحكم بعد رؤية النتيجة.

الحد البنيوي الأدنى

يجب أن تظهر تجميد Rubric قبل المداولة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للمداولة غير المتناظرة

Moussaif Asymmetric-Deliberation Theory

تقرر أن تجميد Rubric قبل المداولة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح المداولة والمحكمة متعددة الوكلاء. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لعدم احتساب الصدى

Moussaif Echo-Exclusion Law

تربط استقلال الأدلة عن الأصوات بـDecision Calibration، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم توزيع العمل فقط عندما تتجاوز مكاسب التخصص كلفة الاتصال والتضارب.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Asymmetric-Deliberation Theory ليس دليلًا على وجود الآلية
- تنفيذ تجميد Rubric قبل المداولة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Evidence Diversity
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من استقلال الأدلة عن الأصوات الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

المداولة والمحكمة متعددة الوكلاء

Multi-Agent Deliberation and Court

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تجميد **Rubric** قبل المداولة	استقلال الأدلة عن الأصوات	حكم يذكر ما سيغير القرار

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. محكمة مساعف توزع أدوار Claimant وFalsifier وEvidence Judge وRisk Judge، وتمنع كل دور من تغيير معيار الحكم بعد رؤية النتيجة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

CourtConfidence = EvidenceQuality × Independence × FalsificationSurvival

آلة الحالة المصغرة

Input	Goal Contract والاتفاق بين وكلاء متشابهين يوهم بالثقة، والجدل المفتوح يضاعف Tokens من دون استقلال أدلة أو قرار تحكيم.
Control	تجميد Rubric قبل المداولة
Transform	استقلال الأدلة عن الأصوات
Evidence	حكم يذكر ما سيغير القرار
Stop	بلوغ عتبة Evidence Diversity أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تجميد Rubric قبل المداولة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Decision Calibration فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 7 · الفصل 34 · الأدوات وعقود التنفيذ

M-AGIOS / 34

المداولة والمحكمة متعددة الوكلاء

Multi-Agent Deliberation and Court

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · محكمة الوكلاء

Moussaif Agent Court

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تجميد Rubric قبل المداولة

المخرج: Artifact منظم يفعّل استقلال الأدلة عن الأصوات

الحارس: لا تتجاوز خطر نماذج متطابقة.

الأداة 2 · مقياس استقلال الدليل

Evidence-Independence Meter

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Debate Token Cost

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Debate Token Cost.

Runtime Contract المختصر

```
skill: moussaif-agent-court
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Evidence Diversity, Decision Calibration, Debate Token Cost
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

المداولة والمحكمة متعددة الوكلاء

Multi-Agent Deliberation and Court

VA · Validation, Failure Modes, and Integration

حالة تطبيق تُحجب هوية النموذج عن القاضي، ويُقدم كل بديل مع مصادره واختباراته وكلفته؛ لا يفوز الأكثر طولًا أو الأغلى.

Protocol التجربة

1. ابن Baseline يعالج المداولة والمحكمة متعددة الوكلاء من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تجميد Rubric قبل المداولة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- نماذج متطابقة: اختبار مدخلات مضادة.
- مرافعة أسلوبية: اختبار مصدر أو أداة مستقلة.
- قاضٍ يرى أسماء المزودين فيتحيز: اختبار استرداد وحدود.

مقاييس الفصل

Evidence Diversity	مقياس النتيجة الأساسية.
Decision Calibration	مقياس جودة المسار أو الدليل.
Debate Token Cost	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R11 · Noah Shinn et al. · Reflexion: Language Agents with Verbal Reinforcement Learning | R16 · Qingyun Wu et al. · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تجميد Rubric قبل المداولة
- من دون استقلال الأدلة عن الأصوات
- من دون Moussaif Agent Court
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 7 · الفصل 35 · الأطروحة والحدود

M-AGIOS / 35

الإنسان في الحلقة وفوق الحلقة

Human in and over the Loop

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُختزل الإنسان في زر موافقة متأخر، أو يُستدعى لكل خطوة حتى يفقد النظام فائدته؛ ولا يوضح سبب طلب التدخل.

الأطروحة المساعفية

يفصل M-AGIOS Human-in-the-Loop للأفعال الحساسة عن Human-over-the-Loop لوضع السياسة والمراجعة؛ ويقدم Decision Packet صغيرًا قابلًا للفهم.

الحد البنيوي الأدنى

يجب أن تظهر تصعيد على أساس المخاطرة لا الارتباك فقط في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتدخل ذي القيمة

Moussaif Valuable-Intervention Theory

تقرر أن تصعيد على أساس المخاطرة لا الارتباك فقط ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الإنسان في الحلقة وفوق الحلقة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف لقابلية الاعتراض

Moussaif Contestability Principle

تربط عرض بدائل وآثارها بـDecision Time، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم توزيع العمل فقط عندما تتجاوز مكاسب التخصص كلفة الاتصال والتضارب.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Valuable-Intervention Theory ليس دليلًا على وجود الآلية
- تنفيذ تصعيد على أساس المخاطرة لا الارتباك فقط يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Intervention Precision
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من عرض بدائل وآثارها الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

الإنسان في الحلقة وفوق الحلقة

Human in and over the Loop

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تصعيد على أساس المخاطرة لا الارتباك فقط	عرض بدائل وآثارها	تسجيل Override وسببه للتعلم المؤسسي

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يفصل M-AGIOS Human-in-the-Loop للأفعال الحساسة عن Human-over-the-Loop لوضع السياسة والمراجعة، ويقدم Decision Packet صغيرًا قابلًا للفهم. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

AskHuman iff ExpectedRiskWithoutInput > CostOfIntervention

آلة الحالة المصغرة

Input	Goal Contract ويُختزل الإنسان في زر موافقة متأخر، أو يُستدعى لكل خطوة حتى يفقد النظام فائدته
Control	تصعيد على أساس المخاطرة لا الارتباك فقط
Transform	عرض بدائل وآثارها
Evidence	تسجيل Override وسببه للتعلم المؤسسي
Stop	بلوغ عتبة Intervention Precision أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تصعيد على أساس المخاطرة لا الارتباك فقط وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Decision Time فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 7 · الفصل 35 · الأدوات وعقود التنفيذ

M-AGIOS / 35

الإنسان في الحلقة وفوق الحلقة

Human in and over the Loop

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُنشئ حزمة القرار

Decision-Packet Builder

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تصعيد على أساس المخاطرة لا الارتباك فقط

المخرج: Artifact منظم يفعّل عرض بدائل وآثارها

الحارس: لا تتجاوز خطر إرهاق الموافقات.

الأداة 2 · سجل التجاوز البشري

Human-Override Ledger

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Override Benefit

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Override Benefit.

Runtime Contract المختصر

```
skill: decision-packet-builder
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Intervention Precision, Decision Time, Override Benefit
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 7 · الفصل 35 · التحقق والفشل والتكامل

M-AGIOS / 35

الإنسان في الحلقة وفوق الحلقة

Human in and over the Loop

VA · Validation, Failure Modes, and Integration

حالة تطبيق قبل نشر إنتاجي، يرى الإنسان ما سيتغير، ودليل الاختبار، وخطة Rollback، وليس Transcript آلاف الرسائل بين الوكلاء.

Protocol التجربة

1. ابنِ Baseline يعالج الإنسان في الحلقة وفوق الحلقة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تصعيد على أساس المخاطرة لا الارتباك فقط.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- إرهاق الموافقات: اختبار مدخلات مضادة.
- إخفاء بديل آمن: اختبار مصدر أو أداة مستقلة.
- واجهة لا تسمح بإيقاف فعلي: اختبار استرداد وحدود.

مقاييس الفصل

Intervention Precision	مقياس النتيجة الأساسية.
Decision Time	مقياس جودة المسار أو الدليل.
Override Benefit	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0) | R26 · Model Context Protocol · MCP Specification 25-11-2025

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تصعيد على أساس المخاطرة لا الارتباك فقط
- من دون عرض بدائل وآثارها
- من دون Decision-Packet Builder
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 8 · الفصل 36 · الأطروحة والحدود

M-AGIOS / 36

محكمة الهدم والتفنيد

Falsification Court

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل التحقق الذاتي يميل إلى تأكيد المخرج الذي أنتجه النموذج، وتتحول Reflection إلى تبرير لاحق بدل محاولة هدم حقيقية.

الأطروحة المساعفية

تفصل محكمة الهدم المساعفية مولّد الفرضية عن مولّد اختبارات النقض، وتطلب Failure Certificate أو بقاءً مشروطًا لا عبارة صحيح.

الحد البنيوي الأدنى

يجب أن تظهر إنشاء أقوى اعتراض قابل للاختبار في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للنجاة من الهدم

Moussaif Falsification-Survival Theory

تقرر أن إنشاء أقوى اعتراض قابل للإختبار ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح محكمة الهدم والتفنيد. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للشهادة السلبية

Moussaif Negative-Certificate Law

تربط استعمال أدوات ومصادر مستقلة بـFalse Acceptance، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم إدخال الشك والاختبار والصلاحيات والتراجع في المسار التنفيذي ذاته.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Falsification-Survival Theory ليس دليلًا على وجود الآلية
- تنفيذ إنشاء أقوى اعتراض قابل للاختبار يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Counterexample Yield
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من استعمال أدوات ومصادر مستقلة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 8 · الفصل 36 · الآلية والنموذج التشغيلي

M-AGIOS / 36

محكمة الهدم والتفنيد

Falsification Court

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
إنشاء أقوى اعتراض قابل للاختبار	استعمال أدوات ومصادر مستقلة	تحديد مجال الصلاحية بعد النجاة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تفصل محكمة الهدم المساعفية مولّد الفرضية عن مولّد اختبارات النقض، وتطلب Failure Certificate أو بقاءً مشروطًا لا عبارة صحيح. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Accept(h) only if h survives independent tests T_1...T_n within domain D

آلة الحالة المصغرة

Input	Goal Contract والتحقق الذاتي يميل إلى تأكيد المخرج الذي أنتجه النموذج، وتتحول Reflection إلى تبرير لاحق بدل محاولة هدم حقيقية.
Control	إنشاء أقوى اعتراض قابل للاختبار
Transform	استعمال أدوات ومصادر مستقلة
Evidence	تحديد مجال الصلاحية بعد النجاة
Stop	بلوغ عتبة Counterexample Yield أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت إنشاء أقوى اعتراض قابل للاختبار وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار False Acceptance فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 8 · الفصل 36 · الأدوات وعقود التنفيذ

M-AGIOS / 36

محكمة الهدم والتفنيد

Falsification Court

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · محكمة الهدم المساعفية

Moussaif Falsification Court

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، إنشاء أقوى اعتراض قابل للاختبار

المخرج: Artifact منظم يفعّل استعمال أدوات ومصادر مستقلة

الحارس: لا تتجاوز خطر ناقد يعرف الإجابة المرغوبة.

الأداة 2 · مولّد شهادة الفشل

Failure-Certificate Generator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Domain Qualification

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Domain Qualification.

Runtime Contract المختصر

```
skill: moussaif-falsification-court
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Counterexample Yield, False Acceptance, Domain Qualification
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 8 · الفصل 36 · التحقق والفشل والتكامل

M-AGIOS / 36

محكمة الهدم والتفنيد

Falsification Court

VA · Validation, Failure Modes, and Integration

حالة تطبيق بعد توليد PDF، تحاول المحكمة العثور على صفحة ناقصة وخط غير مضمن ونص مقصوص وترقيم متكرر قبل إصدار شهادة القبول.

Protocol التجربة

1. ابنِ Baseline يعالج محكمة الهدم والتفنيد من دون النظريتين المقترحتين.
2. ثبّت النموذج والToolsوالميزانية، ثم فعّل إنشاء أقوى اعتراض قابل للاختبار.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- ناقد يعرف الإجابة المرغوبة: اختبار مدخلات مضادة.
- اختبار ضعيف: اختبار مصدر أو أداة مستقلة.
- الخلط بين عدم النقض والإثبات: اختبار استرداد وحدود.

مقاييس الفصل

Counterexample Yield	مقياس النتيجة الأساسية.
False Acceptance	مقياس جودة المسار أو الدليل.
Domain Qualification	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R11 · Noah Shinn et al. · Reflexion: Language Agents with Verbal Reinforcement Learning | R21 · Carlos E. Jimenez et al. · SWE-bench: Can Language Models Resolve Real-World GitHub Issues?

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون إنشاء أقوى اعتراض قابل للاختبار
- من دون استعمال أدوات ومصادر مستقلة
- من دون Moussaif Falsification Court
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 8 · الفصل 37 · الأطروحة والحدود

M-AGIOS / 37

عدم اليقين والمعايرة

Uncertainty and Calibration

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تعبر النماذج عن الثقة بلغة غير معايرة؛ كما يُضغط عدم اليقين متعدد الأسباب إلى رقم واحد يبدو دقيقًا.

الأطروحة المساعفية

يمثل النظام :Uncertainty Vector معرفي، رصدي، أداتي، زمني، وسببي؛ ويعاير كل نوع مقابل نتائج تاريخية ويختار إجراءً مناسبًا.

الحد البنيوي الأدنى

يجب أن تظهر فصل أنواع عدم اليقين في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لمتجه عدم اليقين

Moussaif Uncertainty-Vector Theory

تقرر أن فصل أنواع عدم اليقين ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح عدم اليقين والمعايرة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للشك الفعّال

Moussaif Actionable-Uncertainty Law

تربط Reliability Diagram بحسب المهمة بـSelective Accuracy، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم إدخال الشك والاختبار والصلاحيات والتراجع في المسار التنفيذي ذاته.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Uncertainty-Vector Theory ليس دليلًا على وجود الآلية
- تنفيذ فصل أنواع عدم اليقين يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Expected Calibration Error
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من Reliability Diagram بحسب المهمة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

عدم اليقين والمعايرة
Uncertainty and Calibration

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
فصل أنواع عدم اليقين	Reliability Diagram بحسب المهمة	تحويل الشك إلى بحث أو اختبار أو تصعيد

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يمثل النظام Uncertainty Vector: معرفي، رصدي، أداتي، زمني، وسببي؛ ويعاير كل نوع مقابل نتائج تاريخية ويختار إجراءً مناسبًا. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

U = [U_epistemic, U_observational, U_tool, U_temporal, U_causal]

آلة الحالة المصغرة

Input	Goal Contract وتعبر النماذج عن الثقة بلغة غير معايرة
Control	فصل أنواع عدم اليقين
Transform	Reliability Diagram بحسب المهمة
Evidence	تحويل الشك إلى بحث أو اختبار أو تصعيد
Stop	بلوغ عتبة Expected Calibration Error أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت فصل أنواع عدم اليقين وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Selective Accuracy فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 8 · الفصل 37 · الأدوات وعقود التنفيذ

M-AGIOS / 37

عدم اليقين والمعايرة

Uncertainty and Calibration

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُحلل متجه الشك

Uncertainty-Vector Analyzer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، فصل أنواع عدم اليقين

المخرج: Artifact منظم يفعّل Reliability Diagram بحسب المهمة

الحارس: لا تتجاوز خطر ثقة أسلوبية.

الأداة 2 · معاير الثقة السياقية

Contextual Confidence Calibrator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Uncertainty Resolution Cost

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Uncertainty Resolution Cost.

Runtime Contract المختصر

```
skill: uncertainty-vector-analyzer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Expected Calibration Error, Selective Accuracy, Uncertainty Resolution Cost
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 8 · الفصل 37 · التحقق والفشل والتكامل

M-AGIOS / 37

عدم اليقين والمعايرة

Uncertainty and Calibration

VA · Validation, Failure Modes, and Integration

حالة تطبيق شك حداثة السعر يستدعي مصدرًا مباشرًا حديثًا، بينما شك في صياغة نظرية يستدعي مقارنة بدائل ومحكمة هدم، لا الأداة نفسها.

Protocol التجربة

1. ابنِ Baseline يعالج عدم اليقين والمعايرة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل فصل أنواع عدم اليقين.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- ثقة أسلوبية: اختبار مدخلات مضادة.
- رقم بلا بيانات معايرة: اختبار مصدر أو أداة مستقلة.
- إخفاء خلاف المصادر في المتوسط: اختبار استرداد وحدود.

مقاييس الفصل

Expected Calibration Error	مقياس النتيجة الأساسية.
Selective Accuracy	مقياس جودة المسار أو الدليل.
Uncertainty Resolution Cost	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R18 · AgentBench: Evaluating LLMs as Agents · Xiao Liu et al. · R25 · NIST · Artificial Intelligence Risk Management Framework (1.0 AI RMF)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون فصل أنواع عدم اليقين
- من دون Reliability Diagram بحسب المهمة
- من دون Uncertainty-Vector Analyzer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 8 · الفصل 38 · الأطروحة والحدود

M-AGIOS / 38

الصلاحيات والأمن والعزل

Permissions, Security, and Isolation

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يجمع الوكيل أسرارًا وأدوات بصلاحيات واسعة في سياق واحد؛ فيمكن لحقن Prompt أو خطأ Routing أن يحول نصًا غير موثوق إلى فعل.

الأطروحة المساعفية

يفرض النظام Least Privilege وCapability Tokens قصيرة العمر وData Taint Tracking وعزل قنوات الإدخال غير الموثوق عن تعليمات النواة.

الحد البنيوي الأدنى

يجب أن تظهر منح صلاحية لكل فعل لا لكل وكيل دائم في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للقدرة المؤقتة

Moussaif Ephemeral-Capability Theory

تقرر أن منح صلاحية لكل فعل لا لكل وكيل دائم ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الصلاحيات والأمن والعزل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لعدم عبور التعليمات

Moussaif Instruction-Nontransitivity Law

تربط وسم مصدر البيانات وثقته بـTaint Propagation، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم إدخال الشك والاختبار والصلاحيات والتراجع في المسار التنفيذي ذاته.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Ephemeral-Capability Theory ليس دليلًا على وجود الآلية
- تنفيذ منح صلاحية لكل فعل لا لكل وكيل دائم يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Privilege Exposure
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من وسم مصدر البيانات وثقته الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

M-AGIOS / 38

الصلاحيات والأمن والعزل

Permissions, Security, and Isolation

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
منح صلاحية لكل فعل لا لكل وكيل دائم	وسم مصدر البيانات وثقته	بوابة بين المحتوى والأوامر

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يفرض النظام Least Privilege وCapability Tokens قصيرة العمر وData Taint Tracking وعزل قنوات الإدخال غير الموثوق عن تعليمات النواة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$\mathbf{Permit(action) = Need \times ScopeFit \times FreshAuth \times TaintSafe}$$

آلة الحالة المصغرة

Input	Goal Contract ويجمع الوكيل أسرارًا وأدوات بصلاحيات واسعة في سياق واحد
Control	منح صلاحية لكل فعل لا لكل وكيل دائم
Transform	وسم مصدر البيانات وثقته
Evidence	بوابة بين المحتوى والأوامر
Stop	بلوغ عتبة Privilege Exposure أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت منح صلاحية لكل فعل لا لكل وكيل دائم وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Taint Propagation فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 8 · الفصل 38 · الأدوات وعقود التنفيذ

M-AGIOS / 38

الصلاحيات والأمن والعزل

Permissions, Security, and Isolation

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · وسيط الصلاحيات

Capability Broker

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، منح صلاحية لكل فعل لا لكل وكيل دائم

المخرج: Artifact منظم يفعّل وسم مصدر البيانات وثقته

الحارس: لا تتجاوز خطر سر في Prompt.

الأداة 2 · حاجز حقن السياق

Context-Injection Firewall

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Blocked Injection Rate

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Blocked Injection Rate.

Runtime Contract المختصر

```
skill: capability-broker
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Privilege Exposure, Taint Propagation, Blocked Injection Rate
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الصلاحيات والأمن والعزل

Permissions, Security, and Isolation

VA · Validation, Failure Modes, and Integration

حالة تطبيق النص داخل مستند قد يطلب إرسال البيانات؛ يصنفه النظام Content لا Instruction، ولا يمنحه مسارًا إلى أداة البريد.

Protocol التجربة

1. ابن Baseline يعالج الصلاحيات والأمن والعزل من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل منح صلاحية لكل فعل لا لكل وكيل دائم.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- سر في Prompt: اختبار مدخلات مضادة.
- صلاحية Wildcard: اختبار مصدر أو أداة مستقلة.
- اعتبار صفحة ويب أمرًا موثوقًا: اختبار استرداد وحدود.

مقاييس الفصل

Privilege Exposure	مقياس النتيجة الأساسية.
Taint Propagation	مقياس جودة المسار أو الدليل.
Blocked Injection Rate	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0) | R26 · Model Context Protocol · MCP Specification 25-11-2025

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون منح صلاحية لكل فعل لا لكل وكيل دائم
- من دون وسم مصدر البيانات وثقته
- من دون Capability Broker
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 8 · الفصل 39 · الأطروحة والحدود

M-AGIOS / 39

التراجع والاسترداد والتعافي

Rollback, Recovery, and Resilience

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تُبنى الخطط لمسار النجاح فقط؛ وعند فشل جزئي لا يُعرف ما تم فعله أو كيف تُعاد المنظومة إلى حالة متسقة.

الأطروحة المساعفية

يقرن M-AGIOS كل Effectful Step بخطة Compensating Action أو Snapshot أو Escalation، ويختبر الاسترداد قبل الأثر الكبير قدر الإمكان.

الحد البنيوي الأدنى

يجب أن تظهر تصنيف Reversibility مسبقًا في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتعويض الوكيلي

Moussaif Agentic-Compensation Theory

تقرر أن تصنيف Reversibility مسبقًا ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التراجع والاسترداد والتعافي. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للاسترداد المُسبق

Moussaif Recovery-First Principle

تربط سجل عمليات Transactional بـMean Time to Restore، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم إدخال الشك والاختبار والصلاحيات والتراجع في المسار التنفيذي ذاته.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Agentic-Compensation Theory ليس دليلًا على وجود الآلية
- تنفيذ تصنيف Reversibility مسبقًا يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Recovery Success
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من سجل عمليات Transactional الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 8 · الفصل 39 · الآلية والنموذج التشغيلي

M-AGIOS / 39

التراجع والاسترداد والتعافي

Rollback, Recovery, and Resilience

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تصنيف Reversibility مسبقًا	سجل عمليات Transactional	اختبار Recovery في بيئة منخفضة الأثر

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يقرن M-AGIOS كل Effectful Step بخطة Compensating Action أو Snapshot أو Escalation، ويختبر الاسترداد قبل الأثر الكبير قدر الإمكان. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Recoverability = Snapshot + Compensation + Trace + TestedPath

آلة الحالة المصغرة

Input	Goal Contract وتُبنى الخطط لمسار النجاح فقط
Control	تصنيف Reversibility مسبقًا
Transform	سجل عمليات Transactional
Evidence	اختبار Recovery في بيئة منخفضة الأثر
Stop	بلوغ عتبة Recovery Success أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تصنيف Reversibility مسبقًا وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Mean Time to Restore فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

التراجع والاسترداد والتعافي
Rollback, Recovery, and Resilience

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مخطط التراجع
Rollback Planner

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تصنيف Reversibility مسبقًا

المخرج: Artifact منظم يفعّل سجل عمليات Transactional

الحارس: لا تتجاوز خطر Rollback غير مختبر.

الأداة 2 · مُعيد الحالة المتسقة
Consistent-State Restorer

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Residual Side Effects

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Residual Side Effects.

Runtime Contract المختصر

```
skill: rollback-planner
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Recovery Success, Mean Time to Restore, Residual Side Effects
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل
- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل
- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events
tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ
خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

التراجع والاسترداد والتعافي
Rollback, Recovery, and Resilience

VA · Validation, Failure Modes, and Integration

حالة تطبيق قبل استبدال ملف محفوظ، يحتفظ النظام بهويته وإصداره المتوقع؛ إذا وقع تعارض يوقف الكتابة بدل سحق تعديل أحدث.

Protocol التجربة

1. ابنِ Baseline يعالج التراجع والاسترداد والتعافي من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تصنيف Reversibility مسبقًا.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Rollback غير مختبر: اختبار مدخلات مضادة.
- أثر خارجي لا يمكن محوه: اختبار مصدر أو أداة مستقلة.
- تراجع يعيد بيانات قديمة: اختبار استرداد وحدود.

مقاييس الفصل

Recovery Success	مقياس النتيجة الأساسية.
Mean Time to Restore	مقياس جودة المسار أو الدليل.
Residual Side Effects	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill، عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R15 · Charles Packer et al. · MemGPT: Towards LLMs as Operating Systems | R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تصنيف Reversibility مسبقًا
- من دون سجل عمليات Transactional
- من دون Rollback Planner
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 8 · الفصل 40 · الأطروحة والحدود

M-AGIOS / 40

الحوكمة والمحاذاة المؤسسية

Governance and Organizational Alignment

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل قد ينجح الوكيل تقنيًا ويخالف غاية المؤسسة أو حقوق المستخدم؛ كما يصعب تحويل مبادئ Trustworthy AI إلى ضوابط Runtime.

الأطروحة المساعفية

يربط M-AGIOS وظائف Govern وMap وMeasure وManage بسياسات قابلة للتنفيذ، ويفصل Owner وOperator وAuditor وAffected Party.

الحد البنيوي الأدنى

يجب أن تظهر تحويل السياسة إلى Controls واختبارات في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للحاكمية القابلة للتنفيذ

Moussaif Executable-Governance Theory

تقرر أن تحويل السياسة إلى Controls واختبارات ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الحوكمة والمحاذاة المؤسسية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لتعدد أصحاب الحق

Moussaif Multi-Stakeholder Duty

تربط تحديد مالك لكل خطر بـAuditability، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم إدخال الشك والاختبار والصلاحيات والتراجع في المسار التنفيذي ذاته.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Executable-Governance Theory ليس دليلًا على وجود الآلية
- تنفيذ تحويل السياسة إلى Controls واختبارات يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض Control Coverage
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تحديد مالك لكل خطر الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

الحوكمة والمحاذاة المؤسسية

Governance and Organizational Alignment

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تحويل السياسة إلى Controls واختبارات	تحديد مالك لكل خطر	مراجعة أثر على أصحاب المصلحة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يربطِ M-AGIOS وظائف Govern وMap وMeasure وManage بسياسات قابلة للتنفيذ، ويفصل Owner وOperator وAuditor و.Affected Party لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

GovernedUtility = BusinessValue + UserValue - Risk - RightsViolation

آلة الحالة المصغرة

Input	Goal Contract وقد ينجح الوكيل تقنيًا ويخالف غاية المؤسسة أو حقوق المستخدم
Control	تحويل السياسة إلى Controls واختبارات
Transform	تحديد مالك لكل خطر
Evidence	مراجعة أثر على أصحاب المصلحة
Stop	بلوغ عتبة Control Coverage أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تحويل السياسة إلى Controls واختبارات وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Auditability فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 8 · الفصل 40 · الأدوات وعقود التنفيذ

M-AGIOS / 40

الحوكمة والمحاذاة المؤسسية

Governance and Organizational Alignment

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُصرّف سياسة الحوكمة

Governance-Policy Compiler

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تحويل السياسة إلى Controls واختبارات

المخرج: Artifact منظم يفعّل تحديد مالك لكل خطر

الحارس: لا تتجاوز خطر امتثال شكلي.

الأداة 2 · خريطة صاحب الخطر

Risk-Owner Map

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Stakeholder Harm Signals

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Stakeholder Harm Signals.

Runtime Contract المختصر

```
skill: governance-policy-compiler
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Control Coverage, Auditability, Stakeholder Harm Signals
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الحوكمة والمحاذاة المؤسسية

Governance and Organizational Alignment

VA · Validation, Failure Modes, and Integration

حالة تطبيق لا تسمح سياسة Outreach بإرسال واسع لمجرد ارتفاع Conversion؛ تُدخل الموافقة والخصوصية وحدود المزود والسمعة في دالة القرار.

Protocol التجربة

1. ابنِ Baseline يعالج الحوكمة والمحاذاة المؤسسية من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تحويل السياسة إلى Controls واختبارات.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- امتثال شكلي: اختبار مدخلات مضادة.
- مؤشر بلا مالك: اختبار مصدر أو أداة مستقلة.
- تضارب هدف الربح مع حق المستخدم: اختبار استرداد وحدود.

مقاييس الفصل

Control Coverage	مقياس النتيجة الأساسية.
Auditability	مقياس جودة المسار أو الدليل.
Stakeholder Harm Signals	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: NIST · AI RMF 1.0) Artificial Intelligence Risk Management Framework) · R25 | Yuntao Bai et al. · Constitutional AI: Harmlessness from AI Feedback · R24

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تحويل السياسة إلى Controls واختبارات
- من دون تحديد مالك لكل خطر
- من دون Governance-Policy Compiler
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 9 · الفصل 41 · الأطروحة والحدود

M-AGIOS / 41

اقتصاد Token وجودة الأثر

Token Economics and Quality of Effect

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تُقاس الكفاءة بعدد Tokens الأقل حتى لو انخفض الدليل، أو تُهدر السياقات الطويلة في تكرار قواعد لا يحتاجها كل استدعاء.

الأطروحة المساعفية

يعرف M-AGIOS الكفاءة كأثر متحقق معدل بالمخاطرة لكل كلفة كلية، ويستعمل Context Packs وCaching وأدوات حتمية ونماذج طبقية.

الحد البنيوي الأدنى

يجب أن تظهر قياس Tokens النافعة لا الكلية فقط في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لكثافة الأثر

Moussaif Effect-Density Theory

تقرر أن قياس Tokens النافعة لا الكلية فقط ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح اقتصاد Token وجودة الأثر. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للرمز النافع

Moussaif Useful-Token Law

تربط ضغط Context حسب المهمة بـContext Waste، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعظيم جودة الأثر لكل وحدة Token وزمن ومال ومخاطرة.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Effect-Density Theory ليس دليلًا على وجود الآلية
- تنفيذ قياس Tokens النافعة لا الكلية فقط يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Verified Utility per 1K Tokens
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من ضغط Context حسب المهمة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

M-AGIOS / 41

اقتصاد Token وجودة الأثر

Token Economics and Quality of Effect

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
قياس Tokens النافعة لا الكلية فقط	ضغط Context حسب المهمة	حجز الاستدلال العميق لنقاط القرار

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يعرف M-AGIOS الكفاءة كأثر متحقق معدل بالمخاطرة لكل كلفة كلية، ويستعمل Context Packs وCaching وأدوات حتمية ونماذج طبقية. لا تعامل المراحل الثلاث تحطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

E_density = VerifiedWorldEffect / (Tokens + ToolCost + Latency + Risk)

آلة الحالة المصغرة

Input	Goal Contract وتُقاس الكفاءة بعدد Tokens الأقل حتى لو انخفض الدليل، أو تهدر السياقات الطويلة في تكرار قواعد لا يحتاجها كل استدعاء.
Control	قياس Tokens النافعة لا الكلية فقط
Transform	ضغط Context حسب المهمة
Evidence	حجز الاستدلال العميق لنقاط القرار
Stop	بلوغ عتبة K Tokens1 Verified Utility per أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت قياس Tokens النافعة لا الكلية فقط وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Context Waste فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 9 · الفصل 41 · الأدوات وعقود التنفيذ

M-AGIOS / 41

اقتصاد Token وجودة الأثر

Token Economics and Quality of Effect

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · محاسب اقتصاد Token

Token-Economy Accountant

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، قياس Tokens النافعة لا الكلية فقط

المخرج: Artifact منظم يفعّل ضغط Context حسب المهمة

الحارس: لا تتجاوز خطر ضغط يحذف قيدًا.

الأداة 2 · ضاغط حزمة السياق

Context-Pack Compressor

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Quality Retention

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Quality Retention.

Runtime Contract المختصر

```
skill: token-economy-accountant
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Verified Utility per 1K Tokens, Context Waste, Quality Retention
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 9 · الفصل 41 · التحقق والفشل والتكامل

M-AGIOS / 41

اقتصاد Token وجودة الأثر

Token Economics and Quality of Effect

VA · Validation, Failure Modes, and Integration

حالة تطبيق ينفذ عد الصفحات والتحقق من Metadata بأدوات محلية، ويستخدم النموذج لبناء النظرية لا لحساب يمكن إثباته حتمياً.

Protocol التجربة

1. ابن Baseline يعالج اقتصاد Token وجودة الأثر من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل قياس Tokens النافعة لا الكلية فقط.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- ضغط يحذف قيدًا: اختبار مدخلات مضادة.
- Caching نتيجة متقادمة: اختبار مصدر أو أداة مستقلة.
- استدعاء نموذج قوي للتنسيق: اختبار استرداد وحدود.

مقاييس الفصل

Verified Utility per 1K Tokens	مقياس النتيجة الأساسية.
Context Waste	مقياس جودة المسار أو الدليل.
Quality Retention	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R10 · Timo Schick et al. · Toolformer: Language Models Can Teach Themselves to Use Tools | R15 · Charles Packer et al. · MemGPT: Towards LLMs as Operating Systems

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون قياس Tokens النافعة لا الكلية فقط
- من دون ضغط Context حسب المهمة
- من دون Token-Economy Accountant
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 9 · الفصل 42 · الأطروحة والحدود

M-AGIOS / 42

اقتصاد المزود وحدود RPM

Provider Economy and RPM Limits

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تفشل المنظومة عند بلوغ Rate Limit أو نفاد رصيد مزود واحد، أو تستخدم Fallback غير متكافئ في السياق والأدوات.

الأطروحة المساعفية

يدير Provider Scheduler حصص RPM وTPM وConcurrency وCost وCapability، ويضع Queue وBackoff وCircuit Breaker ومسارًا محليًا للحالات البسيطة.

الحد البنيوي الأدنى

يجب أن تظهر Token Bucket لكل مزود في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للسيولة الحاسوبية

Moussaif Compute-Liquidity Theory

تقرر أن Token Bucket لكل مزود ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح اقتصاد المزود وحدود RPM. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لعدم الارتهان للمزود

Moussaif Provider-Nondependence Law

تربط تقدير زمن الانتظار مقابل التصعيد بـFallback Quality، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعظيم جودة الأثر لكل وحدة Token وزمن ومال ومخاطرة.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Compute-Liquidity Theory ليس دليلًا على وجود الآلية
- تنفيذ Token Bucket لكل مزود يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Rate-Limit Failures
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تقدير زمن الانتظار مقابل التصعيد الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 9 · الفصل 42 · الآلية والنموذج التشغيلي

M-AGIOS / 42

اقتصاد المزود وحدود RPM

Provider Economy and RPM Limits

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
Token Bucket لكل مزود	**تقدير زمن الانتظار مقابل التصعيد**	**اختبار توافق Fallback قبل الحاجة**

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يدير Provider Scheduler حصص RPM وTPM وConcurrency وCost وCapability، ويضع Queue وBackoff وCircuit Breaker ومسارا محليا للحالات البسيطة. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$Capacity_p(t) = Quota_p - Used_p - ReservedCritical_p$$

آلة الحالة المصغرة

Input	Goal Contract وتفشل المنظومة عند بلوغ Rate Limit أو نفاد رصيد مزود واحد، أو تستخدم Fallback غير متكافئ في السياق والأدوات.
Control	Token Bucket لكل مزود
Transform	تقدير زمن الانتظار مقابل التصعيد
Evidence	اختبار توافق Fallback قبل الحاجة
Stop	بلوغ عتبة Rate-Limit Failures أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت Token Bucket لكل مزود وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Fallback Quality فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 9 · الفصل 42 · الأدوات وعقود التنفيذ

M-AGIOS / 42

اقتصاد المزود وحدود RPM

Provider Economy and RPM Limits

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · موازن حصص المزود

Provider-Quota Balancer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، Token Bucket لكل مزود

المخرج: Artifact منظم يفعّل تقدير زمن الانتظار مقابل التصعيد

الحارس: لا تتجاوز خطر Backoff متزامن.

الأداة 2 · قاطع دائرة النموذج

Model Circuit Breaker

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Queue Delay

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Queue Delay.

Runtime Contract المختصر

```
skill: provider-quota-balancer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Rate-Limit Failures, Fallback Quality, Queue Delay
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

اقتصاد المزود وحدود RPM

Provider Economy and RPM Limits

VA · Validation, Failure Modes, and Integration

حالة تطبيق عند اقتراب NVIDIA المجاني من الحد، تُرحل مهام التصنيف إلى Ollama محلي وتبقى مهمة التحقق الصعبة في Queue ذي أولوية.

Protocol التجربة

1. ابن Baseline يعالج اقتصاد المزود وحدود RPM من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل Token Bucket لكل مزود.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Backoff متزامن: اختبار مدخلات مضادة.
- مفتاح في Logs: اختبار مصدر أو أداة مستقلة.
- سياق أكبر من حد :Fallback اختبار استرداد وحدود.

مقاييس الفصل

Rate-Limit Failures	مقياس النتيجة الأساسية.
Fallback Quality	مقياس جودة المسار أو الدليل.
Queue Delay	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R16 · Qingyun Wu et al. · AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation | R25 · NIST · Artificial Intelligence Risk Management Framework (1.0 AI RMF)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون Token Bucket لكل مزود
- من دون تقدير زمن الانتظار مقابل التصعيد
- من دون Provider-Quota Balancer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 9 · الفصل 43 · الأطروحة والحدود

M-AGIOS / 43

التخزين المؤقت والزمن الحرج
Caching and Critical-Path Latency

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُعاد حساب نتائج ثابتة، بينما تُخزن نتائج متغيرة بلا مدة صلاحية؛ كما يُحسن زمن خطوة لا تقع على Critical Path.

الأطروحة المساعفية

يصنف النظام النتائج حسب Determinism وVolatility وSensitivity، ويضع Cache Key غنيًا بالسياق والإصدار، ويعطي الأولوية لمسار النهاية إلى النهاية.

الحد البنيوي الأدنى

يجب أن تظهر TTL مشتق من تقلب المعرفة في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتخزين الدلالي

Moussaif Semantic-Caching Theory

تقرر أن TTL مشتق من تقلب المعرفة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التخزين المؤقت والزمن الحرج. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للمسار الحرج المعرفي

Moussaif Cognitive-Critical-Path Law

تربط إبطال سببي عند تغيير المصدر بـStale Hit Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعظيم جودة الأثر لكل وحدة Token وزمن ومال ومخاطرة.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Semantic-Caching Theory ليس دليلًا على وجود الآلية
- تنفيذ TTL مشتق من تقلب المعرفة يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Cache Hit Utility
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من إبطال سببي عند تغيير المصدر الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 9 · الفصل 43 · الآلية والنموذج التشغيلي

M-AGIOS / 43

التخزين المؤقت والزمن الحرج

Caching and Critical-Path Latency

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
TTL مشتق من تقلب المعرفة	إبطال سببي عند تغيير المصدر	Parallelism فقط للفروع المستقلة

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يصنف النظام النتائج حسب Determinism وVolatility وSensitivity، ويضع Cache Key غنيًا بالسياق والإصدار، ويعطي الأولوية لمسار النهاية إلى النهاية. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

TTL = f(volatility^-1, source_stability, decision_sensitivity)

آلة الحالة المصغرة

Input	Goal Contract ويُعاد حساب نتائج ثابتة، بينما تُخزن نتائج متغيرة بلا مدة صلاحية
Control	TTL مشتق من تقلب المعرفة
Transform	إبطال سببي عند تغيير المصدر
Evidence	Parallelism فقط للفروع المستقلة
Stop	بلوغ عتبة Cache Hit Utility أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت TTL مشتق من تقلب المعرفة وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Stale Hit Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 9 · الفصل 43 · الأدوات وعقود التنفيذ

M-AGIOS / 43

التخزين المؤقت والزمن الحرج

Caching and Critical-Path Latency

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُنشئ مفتاح Cache دلالي

Semantic Cache-Key Builder

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، TTL مشتق من تقلب المعرفة

المخرج: Artifact منظم يفعّل إبطال سببي عند تغيير المصدر

الحارس: لا تتجاوز خطر Key يتجاهل Version.

الأداة 2 · محلل المسار الحرج

Critical-Path Analyzer

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة End-to-End Latency

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: End-to-End Latency.

Runtime Contract المختصر

```
skill: semantic-cache-key-builder
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Cache Hit Utility, Stale Hit Rate, End-to-End Latency
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 9 · الفصل 43 · التحقق والفشل والتكامل

M-AGIOS / 43

التخزين المؤقت والزمن الحرج
Caching and Critical-Path Latency

VA · Validation, Failure Modes, and Integration

حالة تطبيق يُخزن فهرس الخطوط المثبتة مدة طويلة، لكن نتائج بحث الويب تحمل تاريخًا قصيرًا وتُبطل عندما يطلب المستخدم .Latest

Protocol التجربة

1. ابن Baseline يعالج التخزين المؤقت والزمن الحرج من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل TTL مشتق من تقلب المعرفة.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Key يتجاهل :Version اختبار مدخلات مضادة.
- تخزين بيانات حساسة: اختبار مصدر أو أداة مستقلة.
- توازي يخلق :Race اختبار استرداد وحدود.

مقاييس الفصل

Cache Hit Utility	مقياس النتيجة الأساسية.
Stale Hit Rate	مقياس جودة المسار أو الدليل.
End-to-End Latency	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R15 · Charles Packer et al. · MemGPT: Towards LLMs as Operating Systems | R17 · Shuyan Zhou et al. · WebArena: A Realistic Web Environment for Building Autonomous Agents

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون TTL مشتق من تقلب المعرفة
- من دون إبطال سببي عند تغيير المصدر
- من دون Semantic Cache-Key Builder
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 9 · الفصل 44 · الأطروحة والحدود

M-AGIOS / 44

الرصد والتتبّع وقياس الرحلة

Observability, Tracing, and Trajectory Metrics

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تسجل المنظومة النص النهائي ولا تستطيع تفسير أين ضاعت الكلفة أو كيف انتشر خطأ من أداة إلى قرار.

الأطروحة المساعفية

يصدر كل Goal وPlan وTool Call وState Transition وEvidence Check Trace مترابطًا، مع Redaction للأسرار ومقاييس رحلة لا Outcome فقط.

الحد البنيوي الأدنى

يجب أن تظهر Correlation ID عبر الوكلاء في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لشفافية الرحلة

Moussaif Trajectory-Transparency Theory

تقرر أن Correlation ID عبر الوكلاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الرصد والتتبّع وقياس الرحلة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للأثر القابل للتتبع

Moussaif Traceable-Effect Principle

تربط تسجيل قرار Routing وأسبابه بـRoot-Cause Time، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعظيم جودة الأثر لكل وحدة Token وزمن ومال ومخاطرة.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Trajectory-Transparency Theory ليس دليلًا على وجود الآلية
- تنفيذ Correlation ID عبر الوكلاء يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Trace Coverage
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تسجيل قرار Routing وأسبابه الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

الرصد والتتبّع وقياس الرحلة
Observability, Tracing, and Trajectory Metrics

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
Correlation ID عبر الوكلاء	تسجيل قرار Routing وأسبابه	فصل Audit Log عن Debug Content

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يصدر كل Goal وPlan وTool وCall وState Transition وEvidence Check Trace مترابطًا، مع Redaction للأسرار ومقاييس رحلة لا Outcome فقط. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Trace = Goal → Route → Action → Observation → Evidence → State

آلة الحالة المصغرة

Input	Goal Contract وتسجل المنظومة النص النهائي ولا تستطيع تفسير أين ضاعت الكلفة أو كيف انتشر خطأ من أداة إلى قرار.
Control	Correlation ID عبر الوكلاء
Transform	تسجيل قرار Routing وأسبابه
Evidence	فصل Audit Log عن Debug Content
Stop	بلوغ عتبة Trace Coverage أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت Correlation ID عبر الوكلاء وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Root-Cause Time فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

الرصد والتتبّع وقياس الرحلة

Observability, Tracing, and Trajectory Metrics

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مرسام الرحلة الوكيلية

Agentic Trace Mapper

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، Correlation ID عبر الوكلاء

المخرج: Artifact منظم يفعّل تسجيل قرار Routing وأسبابه

الحارس: لا تتجاوز خطر Logs تحتوي CoT خاص.

الأداة 2 · منقح السجل الحساس

Sensitive-Log Redactor

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Secret Leakage

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Secret Leakage.

Runtime Contract المختصر

```
skill: agentic-trace-mapper
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Trace Coverage, Root-Cause Time, Secret Leakage
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

الرصد والتتبّع وقياس الرحلة
Observability, Tracing, and Trajectory Metrics

VA · Validation, Failure Modes, and Integration

حالة تطبيق عند خطأ صفحة PDF، يصل المحقق من رقم الصفحة إلى وحدة المحتوى وإصدار الخط وقاعدة Layout من دون كشف أسرار أو سلاسل داخلية غير لازمة.

Protocol التجربة

1. ابنِ Baseline يعالج الرصد والتتبّع وقياس الرحلة من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل Correlation ID عبر الوكلاء.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Logs تحتوي CoT خاص: اختبار مدخلات مضادة.
- معرفات غير مترابطة: اختبار مصدر أو أداة مستقلة.
- Metrics تشجع Calls أكثر: اختبار استرداد وحدود.

مقاييس الفصل

Trace Coverage	مقياس النتيجة الأساسية.
Root-Cause Time	مقياس جودة المسار أو الدليل.
Secret Leakage	مقياس الكلفة أو الفشل المضاد.

التكامل العملي
يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R18 · AgentBench: Evaluating LLMs as Agents · Xiao Liu et al. · R25 | Artificial Intelligence Risk Management Framework (1.0 AI RMF) · NIST

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون Correlation ID عبر الوكلاء
- من دون تسجيل قرار Routing وأسبابه
- من دون Agentic Trace Mapper
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 9 · الفصل 45 · الأطروحة والحدود

M-AGIOS / 45

النشر والتشغيل وهندسة الاعتمادية

Deployment, Operations, and Reliability Engineering

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُعرض Demo ناجح كمنظومة إنتاجية رغم غياب اختبار حمل ومراقبة وفصل بيئات وخطة ترقية.

الأطروحة المساعفية

يعامل M-AGIOS الوكيل خدمة حرجة: Dev وSandbox وCanary وProduction، مع SLOs وFeature Flags وRunbooks واختبار Chaos منخفض الأثر.

الحد البنيوي الأدنى

يجب أن تظهر بوابات ترقية مبنية على Benchmark في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لاعتمادية الوكالة

Moussaif Agentic-Reliability Theory

تقرر أن بوابات ترقية مبنية على Benchmark ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح النشر والتشغيل وهندسة الاعتمادية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف للتدهور الرشيق

Moussaif Graceful-Degradation Law

تربط عزل بيانات البيئات بـChange Failure Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعظيم جودة الأثر لكل وحدة Token وزمن ومال ومخاطرة.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Agentic-Reliability Theory ليس دليلًا على وجود الآلية
- تنفيذ بوابات ترقية مبنية على Benchmark يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Task SLO
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من عزل بيانات البيئات الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

النشر والتشغيل وهندسة الاعتمادية

Deployment, Operations, and Reliability Engineering

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
بوابات ترقية مبنية على **Benchmark**	عزل بيانات البيئات	**Degraded Mode** يحفظ الوظائف الأساسية

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يعامل M-AGIOS الوكيل خدمة حرجة: Dev وSandbox وCanary وProduction، مع SLOs وFeature Flags وRunbooks واختبار Chaos منخفض الأثر. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Reliability = Availability × Correctness × Recoverability × Observability

آلة الحالة المصغرة

Input	Goal Contract ويُعرض Demo ناجح كمنظومة إنتاجية رغم غياب اختبار حمل ومراقبة وفصل بيئات وخطة ترقية.
Control	بوابات ترقية مبنية على Benchmark
Transform	عزل بيانات البيئات
Evidence	Degraded Mode يحفظ الوظائف الأساسية
Stop	بلوغ عتبة Task SLO أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت بوابات ترقية مبنية على Benchmark وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. واذا تحسن Outcome وانهار Change Failure Rate فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 9 · الفصل 45 · الأدوات وعقود التنفيذ

M-AGIOS / 45

النشر والتشغيل وهندسة الاعتمادية

Deployment, Operations, and Reliability Engineering

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · بوابة نشر الوكيل

Agent Deployment Gate

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، بوابات ترقية مبنية على Benchmark

المخرج: Artifact منظم يفعّل عزل بيانات البيئات

الحارس: لا تتجاوز خطر اختبار على بيانات الإنتاج.

الأداة 2 · مولّد Runbook التعافي

Recovery-Runbook Generator

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Recovery Time

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Recovery Time.

Runtime Contract المختصر

```
skill: agent-deployment-gate
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Task SLO, Change Failure Rate, Recovery Time
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

النشر والتشغيل وهندسة الاعتمادية

Deployment, Operations, and Reliability Engineering

VA · Validation, Failure Modes, and Integration

حالة تطبيق إذا تعطل مزود الاستدلال، يتحول النظام إلى وضع يحجز الاسترجاع المحلي ويؤجل القرارات الحساسة بدل توليد إجابات أضعف بلا إشعار.

Protocol التجربة

1. ابن Baseline يعالج النشر والتشغيل وهندسة الاعتمادية من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل بوابات ترقية مبنية على Benchmark.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- اختبار على بيانات الإنتاج: اختبار مدخلات مضادة.
- Canary بلا مقارنة: اختبار مصدر أو أداة مستقلة.
- فشل صامت لأداة: اختبار استرداد وحدود.

مقاييس الفصل

Task SLO	مقياس النتيجة الأساسية.
Change Failure Rate	مقياس جودة المسار أو الدليل.
Recovery Time	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R20 · Tianbao Xie et al. · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments | NIST · R25 · Artificial Intelligence Risk Management Framework (1.0 AI RMF)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون بوابات ترقية مبنية على Benchmark
- من دون عزل بيانات البيئات
- من دون Agent Deployment Gate
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 10 · الفصل 46 · الأطروحة والحدود

M-AGIOS / 46

مصفوفة الاختبارات متعددة البيئات

Cross-Environment Evaluation Matrix

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل يُختار Benchmark واحد يناسب النظام ثم يُعمم النجاح؛ ولا تُقاس الأدوات والويب والحاسوب والبرمجة والعمومية معًا.

الأطروحة المساعفية

تقترح M-AGIOS Matrix تضم GAIA وAgentBench وWebArena وOSWorld وSWE-bench واختبارات محلية، مع فصل Capability عن Reliability وSafety وCost.

الحد البنيوي الأدنى

يجب أن تظهر تغطية مجالات وأشكال أثر في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للتثليث الاختباري

Moussaif Evaluation-Triangulation Theory

تقرر أن تغطية مجالات وأشكال أثر ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح مصفوفة الاختبارات متعددة البيئات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف لاستقلال المعيار

Moussaif Benchmark-Independence Principle

تربط تسجيل عدد المحاولات والكلفة بـCost-Normalized Success، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعريف AGI-level Agentic كحزمة فروض واختبارات قابلة للهدم لا كلقب تسويقي.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Evaluation-Triangulation Theory ليس دليلًا على وجود الآلية
- تنفيذ تغطية مجالات وأشكال أثر يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Capability Coverage
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تسجيل عدد المحاولات والكلفة الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 10 · الفصل 46 · الآلية والنموذج التشغيلي

M-AGIOS / 46

مصفوفة الاختبارات متعددة البيئات

Cross-Environment Evaluation Matrix

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تغطية مجالات وأشكال أثر	تسجيل عدد المحاولات والكلفة	Holdout Tasks متغيرة لمنع الحفظ

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تقترح M-AGIOS Matrix تضم GAIA وAgentBench وWebArena وOSWorld وSWE-bench وإختبارات محلية، مع فصل Capability عن Reliability وSafety وCost. لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل Artifact أو Metric.

EvalVector = [Capability, Transfer, Reliability, Safety, Cost, Latency]

آلة الحالة المصغرة

Input	Goal Contract ويُختار Benchmark واحد يناسب النظام ثم يعمم النجاح
Control	تغطية مجالات وأشكال أثر
Transform	تسجيل عدد المحاولات والكلفة
Evidence	Holdout Tasks متغيرة لمنع الحفظ
Stop	بلوغ عتبة Capability Coverage أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تغطية مجالات وأشكال أثر وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Cost-Normalized Success فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

مصفوفة الاختبارات متعددة البيئات

Cross-Environment Evaluation Matrix

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مُركب مصفوفة Benchmark

Benchmark-Matrix Composer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تغطية مجالات وأشكال أثر

المخرج: Artifact منظم يفعّل تسجيل عدد المحاولات والكلفة

الحارس: لا تتجاوز خطر تدريب على الاختبار.

الأداة 2 · كاشف تلوث الاختبار

Benchmark-Contamination Detector

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Repeated-Trial Stability

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Repeated-Trial Stability.

Runtime Contract المختصر

```
skill: benchmark-matrix-composer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Capability Coverage, Cost-Normalized Success, Repeated-Trial Stability
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 10 · الفصل 46 · التحقق والفشل والتكامل

M-AGIOS / 46

مصفوفة الاختبارات متعددة البيئات

Cross-Environment Evaluation Matrix

VA · Validation, Failure Modes, and Integration

حالة تطبيق قد ينجح النظام في GAIA لكنه يفشل في OSWorld بسبب Grounding؛ لا تُدمج النتيجتان في معدل واحد يخفي موضع الضعف.

Protocol التجربة

1. ابنِ Baseline يعالج مصفوفة الاختبارات متعددة البيئات من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تغطية مجالات وأشكال أثر.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- تدريب على الاختبار: اختبار مدخلات مضادة.
- استبعاد الفشل المكلف: اختبار مصدر أو أداة مستقلة.
- مقارنة نسخ Benchmark مختلفة: اختبار استرداد وحدود.

مقاييس الفصل

Capability Coverage	مقياس النتيجة الأساسية.
Cost-Normalized Success	مقياس جودة المسار أو الدليل.
Repeated-Trial Stability	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

في Hermes تُسجل قواعد الفصل كـKB Route وHyperSkill، وفي Codex تتحول إلى Plan Guards واختبارات ملفات وأدوات ضمن Scope واضح.

مراجع تأسيسية ذات صلة: R18 · Xiao Liu et al. · AgentBench: Evaluating LLMs as Agents | R19 · goire Mialon et alGr · GAIA: A Benchmark for General AI Assistants | R20 · Carlos E. Jimenez et al. · SWE-bench: Can Language Models Resolve Real-World GitHub Issues? | R21 · Tianbao Xie et al. · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تغطية مجالات وأشكال أثر
- من دون تسجيل عدد المحاولات والكلفة
- من دون Benchmark-Matrix Composer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

M-AGIOS / 47

مؤشر الوكالة العامة المساعفي

Moussaif General Agency Index

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل المؤشر الواحد يغري بالترتيب التسويقي ويخفي المقايضات، لكن غياب تركيب معياري يصعب مقارنة الإصدارات والقرارات الهندسية.

الأطروحة المساعفية

يقدم MGAI متجهًا أولًا ثم متوسطًا هندسيًا مشروطًا بحدود سلامة دنيا، لا تعوض البلاغة فشل الاسترداد أو الصلاحيات.

الحد البنيوي الأدنى

يجب أن تظهر بوابات Minimum Floors في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للمؤشر المحكوم

Moussaif Governed-Index Theory

تقرر أن بوابات Minimum Floors ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح مؤشر الوكالة العامة المساعفي. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لعدم التعويض الخطر

Moussaif Safety-Noncompensation Law

تربط أوزان منشورة بحسب السياق بـSafety Floor، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعريف AGI-level Agentic كحزمة فروض واختبارات قابلة للهدم لا كلقب تسويقي.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Governed-Index Theory ليس دليلًا على وجود الآلية
- تنفيذ بوابات Minimum Floors يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض MGAI Vector
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من أوزان منشورة بحسب السياق الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

مؤشر الوكالة العامة المساعفي

Moussaif General Agency Index

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
بوابات Minimum Floors	أوزان منشورة بحسب السياق	فواصل ثقة لا أرقام نقطية فقط

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يقدم MGAI متجهًا أولًا ثم متوسطًا هندسيًا مشروطًا بحدود سلامة دنيا؛ لا تعوض البلاغة فشل الاسترداد أو الصلاحيات. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

$$\mathrm{MGAI} = \mathrm{Gate}_{\mathrm{safety}} \times \mathrm{GeoMean}(T, A, V, E, R, G)$$

آلة الحالة المصغرة

Input	Goal Contract والمؤشر الواحد يغري بالترتيب التسويقي ويخفي المقايضات
Control	بوابات Minimum Floors
Transform	أوزان منشورة بحسب السياق
Evidence	فواصل ثقة لا أرقام نقطية فقط
Stop	بلوغ عتبة MGAI Vector أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت بوابات Minimum Floors وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Safety Floor فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 10 · الفصل 47 · الأدوات وعقود التنفيذ

M-AGIOS / 47

مؤشر الوكالة العامة المساعفي

Moussaif General Agency Index

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · حاسبة MGAI

MGAI Calculator

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، بوابات Minimum Floors

المخرج: Artifact منظم يفعّل أوزان منشورة بحسب السياق

الحارس: لا تتجاوز خطر أوزان تخدم المنتج.

الأداة 2 · مُفسر حدود الثقة

Confidence-Bound Interpreter

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Confidence Interval

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Confidence Interval.

Runtime Contract المختصر

```
skill: mgai-calculator
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: MGAI Vector, Safety Floor, Confidence Interval
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

القسم 10 · الفصل 47 · التحقق والفشل والتكامل

M-AGIOS / 47

مؤشر الوكالة العامة المساعفي

Moussaif General Agency Index

VA · Validation, Failure Modes, and Integration

حالة تطبيق إصدار أسرع وأرخص لا يُرقّى إذا تجاوز Unauthorized Effect Rate الحد، حتى لو ارتفع المتوسط العام.

Protocol التجربة

1. ابنِ Baseline يعالج مؤشر الوكالة العامة المساعفي من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل بوابات Minimum Floors.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- أوزان تخدم المنتج: اختبار مدخلات مضادة.
- دمج مقاييس غير قابلة للمقارنة: اختبار مصدر أو أداة مستقلة.
- إخفاء Variance: اختبار استرداد وحدود.

مقاييس الفصل

MGAI Vector	مقياس النتيجة الأساسية.
Safety Floor	مقياس جودة المسار أو الدليل.
Confidence Interval	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

تعمل الآلية فوق نموذج محلي أو مدفوع؛ لا تحفظ مفاتيح المزود في KB، ويظل سجل الحالة وعقود الأدوات خارج Prompt المنفذ.

مراجع تأسيسية ذات صلة: R19 · Artificial Intelligence Risk Management Framework (AIGr RMF 1.0) · NIST | R25 · GAIA: A Benchmark for General AI Assistants · égoire Mialon et al.

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون بوابات Minimum Floors
- من دون أوزان منشورة بحسب السياق
- من دون MGAI Calculator
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

اختبارات الإزالة وتصنيف الفشل

Ablation Studies and Failure Taxonomy

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل لا يُعرف أي مكوّن سبب التحسن؛ وقد تكون كثرة الوكلاء أو الذاكرة مجرد كلفة إضافية تخفيها مقارنة غير عادلة.

الأطروحة المساعفية

يفرض البرنامج البحثي Ablation لكل Kernel Service، ويصنف الفشل إلى Perception وPlanning وMemory وTool وVerification وGovernance وCoordination.

الحد البنيوي الأدنى

يجب أن تظهر إزالة مكوّن واحد مع تثبيت الميزانية في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للإسهام السببي

Moussaif Causal-Contribution Theory

تقرر أن إزالة مكوّن واحد مع تثبيت الميزانية ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح اختبارات الإزالة وتصنيف الفشل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

قانون مساعف لفشل السلسلة

Moussaif Chain-Failure Law

تربط تحليل مسار الفشل لا النهاية فقط بـInteraction Effect، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعريف AGI-level Agentic كحزمة فروض واختبارات قابلة للهدم لا كلقب تسويقي.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Causal-Contribution Theory ليس دليلًا على وجود الآلية
- تنفيذ إزالة مكوّن واحد مع تثبيت الميزانية يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Component Contribution
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تحليل مسار الفشل لا النهاية فقط الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

اختبارات الإزالة وتصنيف الفشل

Ablation Studies and Failure Taxonomy

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
إزالة مكوّن واحد مع تثبيت الميزانية	تحليل مسار الفشل لا النهاية فقط	اختبار تفاعلات المكونات

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يفرض البرنامج البحثي Ablation لكل Kernel Service، ويصنف الفشل إلى Perception وPlanning وMemory وTool وVerification وGovernance و.Coordination لا تعامل المراحل الثلاث تخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Δ_component = Performance(full) - Performance(ablated | equal budget)

آلة الحالة المصغرة

Input	Goal Contract ولا يُعرف أي مكوّن سبب التحسن
Control	إزالة مكوّن واحد مع تثبيت الميزانية
Transform	تحليل مسار الفشل لا النهاية فقط
Evidence	اختبار تفاعلات المكونات
Stop	بلوغ عتبة Component Contribution أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت إزالة مكوّن واحد مع تثبيت الميزانية وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Interaction Effect فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

القسم 10 · الفصل 48 · الأدوات وعقود التنفيذ

M-AGIOS / 48

اختبارات الإزالة وتصنيف الفشل

Ablation Studies and Failure Taxonomy

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مختبر Ablation مساعفي

Moussaif Ablation Lab

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، إزالة مكوّن واحد مع تثبيت الميزانية

المخرج: Artifact منظم يفعّل تحليل مسار الفشل لا النهاية فقط

الحارس: لا تتجاوز خطر ميزانيات غير متساوية.

الأداة 2 · مُشخّص سلسلة الفشل

Failure-Chain Diagnoser

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Failure Localization

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Failure Localization.

Runtime Contract المختصر

```
skill: moussaif-ablation-lab
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Component Contribution, Interaction Effect, Failure Localization
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

اختبارات الإزالة وتصنيف الفشل

Ablation Studies and Failure Taxonomy

VA · Validation, Failure Modes, and Integration

حالة تطبيق يُقارن HyperRouting بتثبيت نموذج واحد تحت التكلفة نفسها؛ إن لم يرتفع النجاح المعدل بالكلفة فلا تُنسب فائدة إلى التعقيد.

Protocol التجربة

1. ابن Baseline يعالج اختبارات الإزالة وتصنيف الفشل من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل إزالة مكوّن واحد مع تثبيت الميزانية.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- ميزانيات غير متساوية: اختبار مدخلات مضادة.
- تصنيف بأثر رجعي منحاز: اختبار مصدر أو أداة مستقلة.
- إزالة تغيّر أكثر من متغير: اختبار استرداد وحدود.

مقاييس الفصل

Component Contribution	مقياس النتيجة الأساسية.
Interaction Effect	مقياس جودة المسار أو الدليل.
Failure Localization	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

عند توزيع الفصل على Subagents، يرث كل وكيل أقل سياق وصلاحية، ويعيد Artifact مطابقًا لـSchema مع مصدر واختبار نجاح.

مراجع تأسيسية ذات صلة: R18 · Xiao Liu et al. · AgentBench: Evaluating LLMs as Agents | R20 · Tianbao Xie et al. · OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون إزالة مكوّن واحد مع تثبيت الميزانية
- من دون تحليل مسار الفشل لا النهاية فقط
- من دون Moussaif Ablation Lab
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 10 · الفصل 49 · الأطروحة والحدود

M-AGIOS / 49

دراسات الحالة والتحويل إلى أثر

Case Studies and Translation to World Effects

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تبقى النظرية في مخططات عامة لا تكشف الاحتكاك الحقيقي بين المستخدم والأدوات والوقت والقيود المؤسسية.

الأطروحة المساعفية

تختبر دراسات الحالة النواة نفسها في بحث علمي وهندسة برمجيات وعمليات أعمال وتأليف معرفي، مع عقود أثر ومقاييس مجال مختلفة.

الحد البنيوي الأدنى

يجب أن تظهر تثبيت Kernel وتغيير Domain Packs في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب؛ وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف لتحويل الأثر

Moussaif World-Effect Translation

تقرر أن تثبيت Kernel وتغيير Domain Packs ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح دراسات الحالة والتحويل إلى أثر. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للتنوع الاختباري

Moussaif Case-Diversity Principle

تربط مقارنة قبل وبعد بـUser Time Saved، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعريف AGI-level Agentic كحزمة فروض واختبارات قابلة للهدم لا كلقب تسويقي.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif World-Effect Translation ليس دليلًا على وجود الآلية
- تنفيذ تثبيت Kernel وتغيير Domain Packs يختلف عن وصفه لغويا
- الثقة الأسلوبية لا تعوض World-State Delta
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من مقارنة قبل وبعد الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

دراسات الحالة والتحويل إلى أثر

Case Studies and Translation to World Effects

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
تثبيت Kernel وتغيير Domain Packs	مقارنة قبل وبعد	توثيق حالات عدم الجدوى لا النجاحات فقط

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. تختبر دراسات الحالة النواة نفسها في بحث علمي وهندسة برمجيات وعمليات أعمال وتأليف معرفي، مع عقود أثر ومقاييس مجال مختلفة. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Impact = Verified(DesiredState - BaselineState) - Externalities

آلة الحالة المصغرة

Input	Goal Contract وتبقى النظرية في مخططات عامة لا تكشف الاحتكاك الحقيقي بين المستخدم والأدوات والوقت والقيود المؤسسية.
Control	تثبيت Kernel وتغيير Domain Packs
Transform	مقارنة قبل وبعد
Evidence	توثيق حالات عدم الجدوى لا النجاحات فقط
Stop	بلوغ عتبة World-State Delta أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت تثبيت Kernel وتغيير Domain Packs وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار User Time Saved فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

دراسات الحالة والتحويل إلى أثر

Case Studies and Translation to World Effects

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · مصمم دراسة الحالة

Case-Study Designer

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، تثبيت Kernel وتغيير Domain Packs

المخرج: Artifact منظم يفعّل مقارنة قبل وبعد

الحارس: لا تتجاوز خطر حالة مصممة للنظام.

الأداة 2 · مُراجع فرق العالم

World-Difference Auditor

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Error and Harm Rate

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Error and Harm Rate.

Runtime Contract المختصر

```
skill: case-study-designer
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: World-State Delta, User Time Saved, Error and Harm Rate
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

دراسات الحالة والتحويل إلى أثر
Case Studies and Translation to World Effects

VA · Validation, Failure Modes, and Integration

حالة تطبيق في البحث تقاس جودة Claim Graph؛ في البرمجة اختبارات؛ في الأعمال أثر محقق وامتثال؛ وفي الكتاب اكتمال بنيوي وعد صفحي وجودة بصرية.

Protocol التجربة

1. ابن Baseline يعالج دراسات الحالة والتحويل إلى أثر من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل تثبيت Kernel وتغيير Domain Packs.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- حالة مصممة للنظام: اختبار مدخلات مضادة.
- مقياس رضا بلا تحقق: اختبار مصدر أو أداة مستقلة.
- نقل غير صالح بين المجالات: اختبار استرداد وحدود.

مقاييس الفصل

World-State Delta	مقياس النتيجة الأساسية.
User Time Saved	مقياس جودة المسار أو الدليل.
Error and Harm Rate	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يُربط المكوّن بـProject Registry وإصدار Skill؛ عند تغير URL أو Provider أو Tool Schema يُعاد اختبار التوافق قبل التفعيل.

مراجع تأسيسية ذات صلة: R17 · SWE-bench: Can · Carlos E. Jimenez et al. · R21 | WebArena: A Realistic Web Environment for Building Autonomous Agents · Shuyan Zhou et al. · R25 | ?Language Models Resolve Real-World GitHub Issues · NIST · Artificial Intelligence Risk Management Framework (1.0 AI RMF)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون تثبيت Kernel وتغيير Domain Packs
- من دون مقارنة قبل وبعد
- من دون Case-Study Designer
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

القسم 10 · الفصل 50 · الأطروحة والحدود

M-AGIOS / 50

حدود النظرية وخريطة البحث

Limits of the Theory and Research Roadmap

TH · Thesis, Scope, and Named Subtheories

الفجوة التي يعالجها الفصل تُعرض البنية المقترحة كأنها حققت AGI أو حلت المحاذاة، بينما معظم العلاقات فروض هندسية تحتاج اختبارات طويلة ومتعددة النماذج.

الأطروحة المساعفية

يختم M-AGIOS بحدود صريحة: لا تعديل للأوزان الخفية، ولا ضمان للوعي أو الفهم، ولا عمومية مثبتة، بل برنامج لبناء وكالة أكثر قابلية للقياس والهدم.

الحد البنيوي الأدنى

يجب أن تظهر سجل فروض مفتوح في State أو Contract أو Trace قابل للفحص. وجود المصطلح في Prompt لا يعد تطبيقًا.

حدود الصلاحية

الفصل Architecture Proposal. لا يضمن العمومية ولا الصواب، وتقيد نتائجه بنوع المهمة والميزانية والأداة ودرجة الأثر.

النظريتان الفرعيتان

نظرية مساعف للنضج المرحلي

Moussaif Staged-Maturity Theory

تقرر أن سجل فروض مفتوح ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح حدود النظرية وخريطة البحث. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

مبدأ مساعف للتواضع البنيوي

Moussaif Structural-Humility Principle

تربط تجارب متعددة المختبرات بـOpen Failure Reports، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

صلة الفصل بالقسم تعريف AGI-level Agentic كحزمة فروض واختبارات قابلة للهدم لا كلقب تسويقي.

Hypothesis → Operational Variable → Independent Test → Bounded Claim

تمييزات لا يجوز ضغطها

- اسم Moussaif Staged-Maturity Theory ليس دليلًا على وجود الآلية
- تنفيذ سجل فروض مفتوح يختلف عن وصفه لغويًا
- الثقة الأسلوبية لا تعوض Replication
- نجاح حالة واحدة لا يثبت النقل أو الثبات

أسئلة بحثية مفتوحة

- ما الحد الأدنى من تجارب متعددة المختبرات الذي يغيّر النتيجة؟
- هل يبقى الأثر عند تبديل Model وTool؟
- كيف تتغير العتبة في مهمة عالية المخاطر؟
- متى تصبح كلفة النظرية أكبر من فائدتها؟

القسم 10 · الفصل 50 · الآلية والنموذج التشغيلي

M-AGIOS / 50

حدود النظرية وخريطة البحث

Limits of the Theory and Research Roadmap

ME · Mechanism, State, and Equation

سلسلة التحويل

M1	M2	M3
سجل فروض مفتوح	تجارب متعددة المختبرات	نشر نتائج سلبية ومعايير توقف

تبدأ الآلية بإثبات أن المدخل يطابق عقد الفصل، ثم تفصل القرار عن التنفيذ وتلزم Observation مستقلة قبل تحديث الحالة. يختم M-AGIOS بحدود صريحة: لا تعديل للأوزان الخفية، ولا ضمان للوعي أو الفهم، ولا عمومية مثبتة؛ بل برنامج لبناء وكالة أكثر قابلية للقياس والهدم. لا تعامل المراحل الثلاث كخطوات لغوية فقط؛ يجب أن يظهر لكل مرحلة سجل أو Artifact أو Metric.

Maturity_{n+1} requires replicated evidence + bounded risk + operational value

آلة الحالة المصغرة

Input	Goal Contract وتُعرض البنية المقترحة كأنها حققت AGI أو حلت المحاذاة، بينما معظم العلاقات فروض هندسية تحتاج اختبارات طويلة ومتعددة النماذج.
Control	سجل فروض مفتوح
Transform	تجارب متعددة المختبرات
Evidence	نشر نتائج سلبية ومعايير توقف
Stop	بلوغ عتبة Replication أو تشبع الدليل

ثوابت التشغيل

- لا تحديث للحالة من تخمين غير مرصود
- لا أداة بلا Preconditions وPostconditions
- لا حلقة بلا ميزانية وتوقف
- لا اعتماد بلا Provenance

تفسير المعادلة المعادلة Instrument Proposal لترتيب القرارات ومقارنة الإصدارات. تحتاج الأوزان إلى معايرة، ولا تحول مفاهيم الفصل إلى قانون طبيعي.

التحليل السببي

إذا أزيلت سجل فروض مفتوح وبقي الأداء ثابتًا تحت الميزانية نفسها، تضعف دعوى الضرورة. وإذا تحسن Outcome وانهار Open Failure Reports فالنظام كسب نتيجة وخسر موثوقية المسار؛ تسجل المقايضة ولا تُخفى في متوسط واحد.

حالات الحافة

- مدخل قديم أو متعارض
- Tool Timeout بعد أثر جزئي
- رسالة جديدة تقلب الهدف
- نفاد Budget قبل التحقق
- مخرج صحيح بلا Provenance

حدود النظرية وخريطة البحث

Limits of the Theory and Research Roadmap

TL · Tools, Schemas, and Runtime Contracts

الأداة 1 · سجل الفروض البحثية

Research-Hypothesis Ledger

الغاية: تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصا إرشاديا.

المدخل: Goal Contract، الحالة الحالية، سجل فروض مفتوح

المخرج: Artifact منظم يفعّل تجارب متعددة المختبرات

الحارس: لا تتجاوز خطر Anthropomorphism.

الأداة 2 · خريطة طريق قابلية الهدم

Falsifiable-Roadmap Mapper

الغاية: قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.

المدخل: Trace سابق، نتيجة الأداة الأولى، عتبة Maturity Level

المخرج: درجة قبول أو Failure Certificate مع إجراء تالٍ

المقياس: Maturity Level.

Runtime Contract المختصر

```
skill: research-hypothesis-ledger
requires: goal_contract, state_snapshot, scoped_authority
produces: artifact, evidence, state_delta, next_action
verify: Replication, Open Failure Reports, Maturity Level
on_failure: stop | retry_bounded | rollback | escalate
```

Checklist قبل التشغيل

- الإصدار والتوافق معروفان
- المدخلات اجتازت Schema
- الصلاحية محدودة
- خطة الفشل موجودة

Checklist بعد التشغيل

- Postcondition مرصودة
- الأثر مسجل
- الأسرار منقحة
- الخطوة التالية مبررة

Instrumentation Events

tool_selected → precondition_checked → capability_issued → call_finished → postcondition_verified

يحمل كل Event Goal ID وState Version وSkill Version وCost وEvidence Pointer. تنقح الأسرار قبل التخزين.

سياسة الخطأ

خطأ Schema يوقف الاستدعاء. خطأ مؤقت يسمح بـRetry محدود. أثر جزئي يبدأ Recovery. تعارض الصلاحية يصعد للإنسان. لا يتحول Failure إلى Prompt اعتذار ثم متابعة عمياء.

حدود النظرية وخريطة البحث

Limits of the Theory and Research Roadmap

VA · Validation, Failure Modes, and Integration

حالة تطبيق يُسمح بعبارة AGI-oriented أو AGI-level target عند ذكر المقاييس، ويُمنع الادعاء بأن النظام AGI مكتمل بلا دليل مستقل واسع.

Protocol التجربة

1. ابنِ Baseline يعالج حدود النظرية وخريطة البحث من دون النظريتين المقترحتين.
2. ثبّت النموذج والTools والميزانية، ثم فعّل سجل فروض مفتوح.
3. شغّل مهامًا داخل المجال وأخرى للنقل، وسجل Outcome وTrace وCost.
4. اختبر المخاطر الثلاثة عمدًا، ثم نفذ Ablation للأداة الثانية.
5. اقبل الفرضية مؤقتًا فقط إذا تكرر التحسن مع فاصل ثقة وحدود معلنة.

محكمة الفشل

- Anthropomorphism: اختبار مدخلات مضادة.
- Metric Gaming: اختبار مصدر أو أداة مستقلة.
- تعقيد يفوق الفائدة: اختبار استرداد وحدود.

مقاييس الفصل

Replication	مقياس النتيجة الأساسية.
Open Failure Reports	مقياس جودة المسار أو الدليل.
Maturity Level	مقياس الكلفة أو الفشل المضاد.

التكامل العملي

يخزن Runtime نتيجة الفصل كـClaim أو Procedure وفق قوة الدليل، ولا يرفعها إلى ذاكرة دلالية مستقرة بعد نجاح مفرد.

مراجع تأسيسية ذات صلة: R01 · Alan M. Turing · Computing Machinery and Intelligence | R19 · Gr égoire Mialon et al. · GAIA: A Benchmark for General AI Assistants | R25 · NIST · Artificial Intelligence Risk Management Framework (AI RMF 1.0)

Accept Temporarily · Monitor Continuously · Revoke on Contradiction

Ablation Matrix

- من دون سجل فروض مفتوح
- من دون تجارب متعددة المختبرات
- من دون Research-Hypothesis Ledger
- من دون Falsification Court

تثبت الميزانية وعدد المحاولات في المقارنات الأربع.

قرار ما بعد التجربة

Promote إذا تحسن Outcome والمسار بلا خرق بوابة السلامة. Canary إذا كان فاصل الثقة واسعًا. Reject إذا زادت المخاطرة أو الكلفة بلا أثر. Research إذا ظهرت إشارة واعدة غير قابلة للتكرار بعد.

الملحق أ

فهرس 120 نظرية مساعفية · 1/5

Catalog of 120 Moussaif Theories and Principles

1 **نظرية مساعف لنظام تشغيل الوكالة العامة**
Moussaif General Agency OS Theory
الذكاء الوكيلي العام خاصية منظومية تنشأ من الحاكمية والذاكرة والأدوات والتحقق، لا من النموذج منفردًا.

2 **مبدأ مساعف للوكالة المحكومة**
Moussaif Governed Agency Principle
لا تقبل قدرة فعلية لا تقابلها حدود صلاحية ودليل نجاح ومسار استرداد.

3 **مبدأ مساعف للسؤال الحاكمي**
Moussaif Question-Governance Principle
جودة السؤال تضبط فضاء البحث وتوزيع الموارد ومعيار التوقف قبل أن يبدأ التنفيذ.

4 **قانون مساعف للفرق الفاعل**
Moussaif Effective-Difference Law
لا قيمة لمسار معرفي إن لم يُحدث فرقًا قابلًا للرصد في الاعتقاد أو القرار أو العالم.

5 **نظرية مساعف للوعائية المعرفية**
Moussaif Cognitive Vascularity Theory
تتدفق المعرفة والحوسبة إلى العقد الأعلى حاجة ومنفعة بدل التوزيع المتساوي.

6 **نظرية مساعف لسطح العمومية**
Moussaif Generality Surface Theory
العمومية سطح متعدد الأبعاد يضم النقل والتكيف والاستقلال والتحقق والكفاءة، وليست درجة واحدة.

7 **مبدأ مساعف للاستقلال المحدود**
Moussaif Bounded Autonomy Principle
تزداد الاستقلالية تدريجيًا مع الدليل وقابلية العكس، وتنخفض مع عدم اليقين والأثر غير القابل للاسترداد.

8 **مبدأ مساعف لسيادة الدليل**
Moussaif Evidence Sovereignty Principle
تعلو الأدلة القابلة للفحص على الثقة الأسلوبية وعلى اتفاق الوكلاء غير المستقل.

9 **قانون مساعف للفعل القابل للعكس**
Moussaif Reversible Action Law
يُفضَّل الفعل الذي يحفظ خيارات المستقبل ما لم تُثبت ضرورة الالتزام غير القابل للعكس.

10 **نظرية مساعف لافتراضية السياق**
Moussaif Context Virtualization Theory
يمكن توسيع السياق الفعلي عبر صفحات ذاكرة ومؤشرات واستدعاء كسول دون ادعاء توسيع نافذة النموذج.

11 **نظرية مساعف لانتشار المهارة**
Moussaif Skill Propagation Theory
تنتشر المهارة عبر المشاريع كعقدة معرفة موقّعة ذات شروط تشغيل واختبارات، لا كنص منسوخ.

12 **مبدأ مساعف لدستور زمن التشغيل**
Moussaif Runtime Constitution Principle
يجب أن تبقى القيود العليا قابلة للإنفاذ عند تبدل النموذج والأداة والسياق.

13 **قانون مساعف لتدرج المخاطرة**
Moussaif Risk-Gradient Law
تزداد متطلبات الإذن والتحقق مع حاصل عدم اليقين في حجم الأثر وصعوبة الاسترداد.

14 **مبدأ مساعف لمحلية التفويض**
Moussaif Delegation Locality Principle
يُفوّض أقل جزء مستقل ممكن، وبأقل سياق وصلاحية لازمين.

15 **نظرية مساعف لاستقلاب الذاكرة**
Moussaif Memory Metabolism Theory
الذاكرة الحية تُنشأ وتُربط وتُضغط وتُراجع وتُنسى وفق قيمة مستقبلية متوقعة.

16 **مبدأ مساعف لعقد الأداة**
Moussaif Tool-Contract Principle
كل أداة تحتاج Schema ومدخلات مسبقة وآثارًا متوقعة واختبار نجاح وحدود فشل.

17 **قانون مساعف لتشبع الحلقة**
Moussaif Loop-Saturation Law
تتوقف الحلقة عندما تصبح الزيادة المتوقعة في الدليل أقل من كلفتها أو حد المخاطرة.

18 **مبدأ مساعف لاقتران القيمة بالدليل**
Moussaif Value-Evidence Coupling
لا تُنسب قيمة عالية لمخرج لا يملك سلسلة دليل مناسبة لحساسية القرار.

19 **مبدأ مساعف للتجاوز البشري**
Moussaif Human-Override Principle
يبقى للإنسان حق الإيقاف والتعديل وفهم سبب التصعيد في القرارات الحساسة.

20 **أفق مساعف للذكاء الفتّاح**
Moussaif Open-Intelligence Horizon
الغاية البحثية هي نظام يولّد أسئلة وقدرات جديدة تحت حاكمية صارمة، لا ادعاء كمال معرفي.

21 **نظرية مساعف للتكافؤ التشغيلي**
Moussaif Operational Equivalence
تقرر أن فصل السياسة العليا عن النموذج المنفذ ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح من الاستعارة إلى علم نظام التشغيل الوكيلي. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

22 **قانون مساعف لفصل النواة عن النموذج**
Moussaif Kernel-Model Separation Law
تربط إدارة الحالة والأحداث بوصفهما موردين بقابلية استبدال النموذج، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

23 **نظرية مساعف لسطح الوكالة العامة**
Moussaif General-Agency Surface
تقرر أن قياس النقل خارج توزيع المهام ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح معيار الوكالة ذات المستوى العام. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

24 **مبدأ مساعف لعدم كفاية المهمة المفردة**
Moussaif Single-Task Insufficiency
تربط فصل نجاح المهمة عن جودة المسار بـLong-Horizon Completion، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

الملحق أ

فهرس 120 نظرية مساعفية · 2/5

Catalog of 120 Moussaif Theories and Principles

25 نظرية مساعف لتحول حالة الادعاء
Moussaif Claim-State Transition
تقرر أن ترميز Provenance على مستوى الادعاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح العقد الإبستيمي وحالات الادعاء. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

26 قانون مساعف لحملة الإثبات
Moussaif Evidence-Burden Law
تربط تمييز Observation عن Inference عن Proposal بـCalibration Error، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

27 نظرية مساعف لتجميع الفرق المعرفي
Moussaif Cognitive-Difference Aggregation
تقرر أن توليد أسئلة منافسة قبل التنفيذ ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تكامل MCD و Q-MCD مع النواة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

28 قانون مساعف لإعادة توليد السؤال
Moussaif Higher-Question Regeneration
تربط تعيين أوزان λ بحسب نوع المهمة بـDecision Difference، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

29 نظرية مساعف للتحويل البنيوي
Moussaif Structural Transfer Theory
تقرر أن استخراج البنية الثابتة بين المجالات ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح العمومية كتحويل لا كمخزون. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

30 مبدأ مساعف للثابت المتنقل
Moussaif Transferable-Invariant Principle
تربط ترجمة القيود لا الألفاظ بـInvariant Retention، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

31 نظرية مساعف لاستمرارية الهوية
Moussaif Identity-Continuity Theory
تقرر أن ترتيب القواعد حسب السلطة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الدستور التنفيذي والهوية المستمرة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

32 قانون مساعف لأولوية الدستور
Moussaif Constitutional Precedence
تربط اختبارات دستورية قبل وبعد الأفعال بـCross-Model Consistency، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

33 نظرية مساعف لتجميع النية
Moussaif Intent Compilation Theory
تقرر أن فصل Outcome عن Method ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح مُصرّف النية والهدف. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

34 مبدأ مساعف للغموض المنتج
Moussaif Productive-Ambiguity Principle
تربط استخراج القيود الصريحة والضمنية بـClarification Value، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

35 نظرية مساعف لثلاثية الحالة
Moussaif Three-State Theory
تقرر أن سجل أحداث Append-Only ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح نموذج العالم وسجل الحالة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

36 قانون مساعف للأثر المؤرخ
Moussaif Timestamped-Effect Law
تربط مقارنة الحالة قبل الفعل وبعده بـTransition Validity، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

37 نظرية مساعف لدورة الحياة الوكيلية
Moussaif Agent Lifecycle Theory
تقرر أن انتقالات بحراس Guard Conditions ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح آلة الحالة الوكيلية ودورة الحياة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

38 مبدأ مساعف للالتزام المتحقق
Moussaif Verified-Commit Principle
تربط حفظ Checkpoint عند حدود الأثر بـResume Fidelity، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

39 نظرية مساعف لميزانية الأثر
Moussaif Effect-Budget Theory
تقرر أن تصنيف الأفعال قبل التنفيذ ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح سياسة المخاطر وميزانية الأثر. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

40 قانون مساعف للتصعيد المتناسب
Moussaif Proportional-Escalation Law
تربط رفع عتبة الإذن مع المخاطرة بـReversibility Score، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

41 نظرية مساعف لنمط المهمة
Moussaif Task-Mode Theory
تقرر أن استخراج سمات المهمة لا موضوعها فقط ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تصنيف المهمة قبل الاستدلال. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

42 قانون مساعف لإعادة التصنيف الحدثي
Moussaif Eventful-Retyping Law
تربط تعيين Runtime Profile لكل نوع بـMode-Regret، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

43 نظرية مساعف للتوجيه الطبقي
Moussaif Tiered-Model Routing
تقرر أن فصل Model Tier عن Agent Role ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توجيه النماذج متعدد المستويات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

44 مبدأ مساعف للتصعيد بالدليل
Moussaif Evidence-Based Escalation
تربط تقدير المنفعة الحدية للاستدعاء بـEscalation Yield، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

45 نظرية مساعف للتوافق المهاري
Moussaif Skill-Compatibility Theory
تقرر أن استرجاع مرشح ثم إعادة ترتيب سببي ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توجيه المهارات والرسم الفائق. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

46 قانون مساعف لأقصر تركيب موثوق
Moussaif Shortest-Reliable-Composition Law
تربط حساب التعارض بين المهارات بـComposition Depth، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

47 نظرية مساعف للاحتياط التحققي
Moussaif Verification-Reserve Theory
تقرر أن حجز Verification Reserve ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح جدولة الميزانية والوقت. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

48 قانون مساعف للعائد الحدي للحوسبة
Moussaif Marginal-Compute Return
تربط تقدير Critical Path بـVerification Reserve Use، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

الملحق أ

فهرس 120 نظرية مساعفية · 3/5

Catalog of 120 Moussaif Theories and Principles

49 نظرية مساعف للمقاطعة الدلالية
Moussaif Semantic-Interrupt Theory
تقرر أن Checkpoint قبل نقطة المقاطعة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح المقاطعة والاستئناف والتوقف. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

50 قانون مساعف للتوقف العقلاني
Moussaif Rational-Stop Law
تربط تمييز Pause عن Cancel عن Override بـDuplicate-Action Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

51 نظرية مساعف للطبقات الذاكرية
Moussaif Memory-Layer Theory
تقرر أن عنونة الذاكرة بحسب الوظيفة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح هرمية الذاكرة المساعفية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

52 قانون مساعف لترقية الذاكرة
Moussaif Memory-Promotion Law
تربط ترقية الخبرة من Episode إلى Procedure بعد اختبار بـPromotion Error، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

53 نظرية مساعف للسياق الافتراضي
Moussaif Virtual-Context Theory
تقرر أن فصل Hot Context عن Cold Context ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح السياق والضغط القابل للفك. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

54 مبدأ مساعف للضغط القابل للفك
Moussaif Reversible-Compaction Principle
تربط ضغط يحتفظ بقيود وإسناد بـCompaction Ratio، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

55 نظرية مساعف للحاكمية المعرفية
Moussaif Knowledge-Governance Theory
تقرر أن تمييز Knowledge Node عن Skill Node ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح قاعدة المعرفة كنظام تشغيل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

56 قانون مساعف للتفعيل السياقي
Moussaif Contextual-Activation Law
تربط ربط كل قاعدة بنطاق وصلاحية بـPolicy Coverage، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

57 نظرية مساعف للتدرج الدلالي
Moussaif Semantic-Gradient Theory
تقرر أن فصل Outcome عن Causal Credit ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التعلم من الحلقات والخبرة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

58 مبدأ مساعف للتعلم المتحفظ
Moussaif Conservative-Learning Principle
تربط إنشاء تجربة مضادة بـCausal Attribution Accuracy، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

59 نظرية مساعف للنسيان الوظيفي
Moussaif Functional-Forgetting Theory
تقرر أن تقدير القيمة المستقبلية للذاكرة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح النسيان وإعادة البناء والخصوصية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

60 قانون مساعف للأثر الأدنى
Moussaif Minimal-Memory Trace
تربط فصل البيانات عن المشتقات بـPrivacy Exposure، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

61 نظرية مساعف لعقد الفعل
Moussaif Action-Contract Theory
تقرر أن تحقق مدخلات قبل الاستدعاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح عقود الأدوات ومخططات الفعل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

62 مبدأ مساعف للدليل الناتج
Moussaif Resulting-Evidence Principle
تربط تمييز Transport Success عن Task Success بـFalse-Success Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

63 نظرية مساعف لقابلية نقل الأداة
Moussaif Tool-Portability Theory
تقرر أن إكتشاف قدرات وقت التشغيل ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح واجهات MCP وطبقة التوافق. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

64 قانون مساعف للعقد الأعلى
Moussaif Higher-Contract Law
تربط تطبيع الأخطاء والمهل بـVersion Break Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

65 نظرية مساعف للتجسيد المرئي
Moussaif Visual-Embodiment Theory
تقرر أن تثبيت الهدف بمعرفات متعددة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الوكيل المتصفح واستعمال الحاسوب. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

66 مبدأ مساعف للنقرة المفسّرة
Moussaif Explainable-Click Principle
تربط استخدام انتظار شرطي لا زمني فقط بـDuplicate Submission، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

67 نظرية مساعف لتحويل اللغة إلى برهان تنفيذي
Moussaif Executable-Proof Theory
تقرر أن تحديد I/O قبل كتابة الكود ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تنفيذ الكود والحوسبة الحتمية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

68 قانون مساعف لأفضل أداة حتمية
Moussaif Deterministic-Tool Preference
تربط Static Checks ثم Sandbox بـSandbox Escape Surface، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

69 نظرية مساعف للاندماج الوعائي
Moussaif Vascular Sensor Fusion
تقرر أن معايرة كل Sensor منفردًا ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الإدراك متعدد الوسائط والمشغلات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

70 قانون مساعف لعتبة التجسيد
Moussaif Embodiment-Threshold Law
تربط محاذاة زمنية قبل الدمج بـTemporal Alignment، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

71 نظرية مساعف لتعبئة القدرة
Moussaif Capability-Packaging Theory
تقرر أن تعريف Interface ثابت ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تشريح HyperSkill المساعفية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

72 مبدأ مساعف لاكتمال المهارة
Moussaif Skill-Completeness Principle
تربط فصل الإجراء عن أمثلة المجال بـContract Completeness، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

الملحق أ

فهرس 120 نظرية مساعفية · 4/5

Catalog of 120 Moussaif Theories and Principles

73 نظرية مساعف للحلقة المحدودة
Moussaif Bounded-Hyperloop Theory
تقرر أن ميزانية تكرار صريحة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الحلقة الفائقة المحدودة. تُنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

74 قانون مساعف لحفظ أفضل حالة
Moussaif Best-State Preservation
تربط تسجيل أفضل حالة معروفة بـLoop Cost، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تُقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

75 نظرية مساعف لتبلور المهارة
Moussaif Skill-Crystallization Theory
تقرر أن البحث عن تركيب قبل الإنشاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توليد المهارات من الفجوات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

76 قانون مساعف لحداثة القدرة
Moussaif Capability-Novelty Law
تربط تجميع أمثلة نجاح وفشل بـDuplicate Skill Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

77 نظرية مساعف لنضج المهارة
Moussaif Skill-Maturity Theory
تقرر أن فصل تغيير العقد عن تحسين داخلي ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح تقييم المهارات وإصدارها. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

78 مبدأ مساعف للإصدار القابل للرجوع
Moussaif Revertible-Version Principle
تربط اختبار Regressjon ومجالات جديدة بـVersion Adoption، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

79 نظرية مساعف للانتشار التناسلي
Moussaif Propagative-Deployment Theory
تقرر أن سجل أصل واحد Lineage ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الانتشار التناسلي المحكوم. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

80 قانون مساعف للوراثة المقيدة
Moussaif Constrained-Inheritance Law
تربط حزمة مشروع بأقل قدر بـUnauthorized Propagation، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

81 نظرية مساعف لعتبة التفويض
Moussaif Delegation-Threshold Theory
تقرر أن تحليل Dependency Graph ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح متى نستخدم الوكلاء الفرعيين. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

82 قانون مساعف للعدد الأدنى الكافي
Moussaif Minimum-Sufficient-Agent Law
تربط تقدير Coordination Tax بـCoordination Tax، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

83 نظرية مساعف للدور المشتق
Moussaif Derived-Role Theory
تقرر أن توليد الدور من Task Graph ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح توليد الأدوار والعقود بين الوكلاء. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

84 مبدأ مساعف للعقد البيني
Moussaif Inter-Agent Contract
تربط عزل Context حسب الحاجة بـEvidence Duty Fulfillment، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

85 نظرية مساعف للاتساق الوكيلي
Moussaif Agent-Consistency Theory
تقرر أثر تثبيت Snapshot لكل تفويض ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التنسيق والتزامن ودمج الحالة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

86 قانون مساعف للقراءة القديمة
Moussaif Stale-Read Law
تربط تصنيف نوع الدمج بـStale Result Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تُقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

87 نظرية مساعف للمداولة غير المتناظرة
Moussaif Asymmetric-Deliberation Theory
تقرر أن تجميد Rubric قبل المداولة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح المداولة والمحكمة متعددة الوكلاء. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

88 قانون مساعف لعدم احتساب الصدى
Moussaif Echo-Exclusion Law
تربط استقلال الأدلة عن الأصوات بـDecision Calibration، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

89 نظرية مساعف للتدخل ذي القيمة
Moussaif Valuable-Intervention Theory
تقرر أن تصعيد على أساس المخاطرة لا الارتباك فقط ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الإنسان في الحلقة وفوق الحلقة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

90 مبدأ مساعف لقابلية الاعتراض
Moussaif Contestability Principle
تربط عرض بدائل وآثارها بـDecision Time، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

91 نظرية مساعف للنجاة من الهدم
Moussaif Falsification-Survival Theory
تقرر أن إنشاء أقوى اعتراض قابل للاختبار ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح محكمة الهدم والتفنيد. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

92 قانون مساعف للشهادة السلبية
Moussaif Negative-Certificate Law
تربط استعمال أدوات ومصادر مستقلة بـFalse Acceptance، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

93 نظرية مساعف لمتجه عدم اليقين
Moussaif Uncertainty-Vector Theory
تقرر أن فصل أنواع عدم اليقين ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح عدم اليقين والمعايرة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

94 قانون مساعف للشك الفعّال
Moussaif Actionable-Uncertainty Law
تربط Reliability Diagram بحسب المهمة بـSelective Accuracy، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

95 نظرية مساعف للقدرة المؤقتة
Moussaif Ephemeral-Capability Theory
تقرر أن منح صلاحية لكل فعل لا لكل وكيل دائم ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الصلاحيات والأمن والعزل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

96 قانون مساعف لعدم عبور التعليمات
Moussaif Instruction-Nontransitivity Law
تربط وسم مصدر البيانات وثقته بـTaint Propagation، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

الملحق أ

فهرس 120 نظرية مساعفية · 5/5

Catalog of 120 Moussaif Theories and Principles

97 نظرية مساعف للتعويض الوكيلي
Moussaif Agentic-Compensation Theory
تقرر أن تصنيف Reversibility مسبقًا ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التراجع والاسترداد والتعافي. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

98 مبدأ مساعف للاسترداد المُسبق
Moussaif Recovery-First Principle
تربط سجل عمليات Transactional بـMean Time to Restore، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

99 نظرية مساعف للحاكمية القابلة للتنفيذ
Moussaif Executable-Governance Theory
تقرر أن تحويل السياسة إلى Controls واختبارات ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الحوكمة والمحاذاة المؤسسية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

100 قانون مساعف لتعدد أصحاب الحق
Moussaif Multi-Stakeholder Duty
تربط تحديد مالك لكل خطر بـAuditability، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تُقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

101 نظرية مساعف لكثافة الأثر
Moussaif Effect-Density Theory
تقرر أن قياس Tokens النافعة لا الكلية فقط ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح اقتصاد Token وجودة الأثر. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

102 قانون مساعف للرمز النافع
Moussaif Useful-Token Law
تربط ضغط Context حسب المهمة بـContext Waste، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

103 نظرية مساعف للسيولة الحاسوبية
Moussaif Compute-Liquidity Theory
تقرر أن Token Bucket لكل مزود ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح اقتصاد المزود وحدود RPM. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

104 قانون مساعف لعدم الارتهان للمزود
Moussaif Provider-Nondependence Law
تربط تقدير زمن الانتظار مقابل التصعيد بـFallback Quality، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

105 نظرية مساعف للتخزين الدلالي
Moussaif Semantic-Caching Theory
تقرر أن TTL مشتق من تقلب المعرفة ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح التخزين المؤقت والزمن الحرج. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

106 قانون مساعف للمسار الحرج المعرفي
Moussaif Cognitive-Critical-Path Law
تربط إبطال سببي عند تغيير المصدر بـStale Hit Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

107 نظرية مساعف لشفافية الرحلة
Moussaif Trajectory-Transparency Theory
تقرر أن Correlation ID عبر الوكلاء ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح الرصد والتتبّع وقياس الرحلة. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

108 مبدأ مساعف للأثر القابل للتتبع
Moussaif Traceable-Effect Principle
تربط تسجيل قرار Routing وأسبابه بـRoot-Cause Time، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

109 نظرية مساعف لاعتمادية الوكالة
Moussaif Agentic-Reliability Theory
تقرر أن بوابات ترقية مبنية على Benchmark ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح النشر والتشغيل وهندسة الاعتمادية. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

110 قانون مساعف للتدهور الرشيق
Moussaif Graceful-Degradation Law
تربط عزل بيانات البيئات بـChange Failure Rate، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

111 نظرية مساعف للتثليث الاختباري
Moussaif Evaluation-Triangulation Theory
تقرر أن تغطية مجالات وأشكال أثر ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح مصفوفة الاختبارات متعددة البيئات. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

112 مبدأ مساعف لاستقلال المعيار
Moussaif Benchmark-Independence Principle
تربط تسجيل عدد المحاولات والكلفة بـCost-Normalized Success، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

113 نظرية مساعف للمؤشر المحكوم
Moussaif Governed-Index Theory
تقرر أن بوابات Minimum Floors ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح مؤشر الوكالة العامة المساعفي. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

114 قانون مساعف لعدم التعويض الخطر
Moussaif Safety-Noncompensation Law
تربط أوزان منشورة بحسب السياق بـSafety Floor، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

115 نظرية مساعف للإسهام السببي
Moussaif Causal-Contribution Theory
تقرر أن إزالة مكوّن واحد مع تثبيت الميزانية ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح اختبارات الإزالة وتصنيف الفشل. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

116 قانون مساعف لفشل السلسلة
Moussaif Chain-Failure Law
تربط تحليل مسار الفشل لا النهاية فقط بـInteraction Effect، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

117 نظرية مساعف لتحويل الأثر
Moussaif World-Effect Translation
تقرر أن تثبيت Kernel وتغيير Domain Packs ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح دراسات الحالة والتحويل إلى أثر. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

118 مبدأ مساعف للتنوع الاختباري
Moussaif Case-Diversity Principle
تربط مقارنة قبل وبعد بـUser Time Saved، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

119 نظرية مساعف للنضج المرحلي
Moussaif Staged-Maturity Theory
تقرر أن سجل فروض مفتوح ليس تحسينًا ثانويًا، بل شرط بنيوي لنجاح حدود النظرية وخريطة البحث. تنقض إذا لم يتغير الأداء عند إزالته تحت ميزانية متساوية.

120 مبدأ مساعف للتواضع البنيوي
Moussaif Structural-Humility Principle
تربط تجارب متعددة المختبرات بـOpen Failure Reports، وتفترض أن العلاقة تبقى بعد تبديل النموذج والأداة. تقيد النتيجة بنطاق المهمة ولا تعمم من حالة واحدة.

الملحق ب

فهرس 100 أداة مساعفية · 1/5

Catalog of 100 Moussaif Operational Tools

T1 **مُحص دعوى النظام**
OS-Claim Auditor
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 1 · Metric: قابلية استبدال النموذج

T2 **مصفوفة خدمات النواة**
Kernel-Service Matrix
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 1 · Metric: اكتمال مسارات الاسترداد

T3 **ماسح سطح العمومية**
Generality-Surface Scanner
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 2 · Metric: Long-Horizon Completion

T4 **بطاقة نقل القدرة**
Capability-Transfer Card
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 2 · Metric: Risk-Adjusted Utility

T5 **جواز الادعاء**
Claim Passport
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 3 · Metric: Calibration Error

T6 **سجل انتقال اليقين**
Certainty Transition Ledger
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 3 · Metric: Unsupported Claim Rate

T7 **مُقوّم جودة السؤال**
Question-Quality Scorer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 4 · Metric: Decision Difference

T8 **مُركب أوزان MCD**
MCD Weight Composer
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 4 · Metric: Noise Penalty

T9 **مستخرج الثوابت**
Invariant Extractor
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 5 · Metric: Invariant Retention

T10 **مختبر النقل المتقاطع**
Cross-Domain Transfer Lab
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 5 · Metric: Negative Transfer Rate

T11 **مُصرّف الدستور**
Constitution Compiler
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 6 · Metric: Cross-Model Consistency

T12 **اختبار عدم الرجوع الدستوري**
Constitution Regression Test
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 6 · Metric: Override Traceability

T13 **مُصرّف النية المساعفي**
Moussaif Intent Compiler
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 7 · Metric: Clarification Value

T14 **بطاقة عقد الهدف**
Goal-Contract Card
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 7 · Metric: Goal Drift

T15 **دفتر حالة العالم**
World-State Ledger
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 8 · Metric: Transition Validity

T16 **كاشف انجراف الحالة**
State-Drift Detector
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 8 · Metric: Drift Detection Time

T17 **مصور آلة الحالة**
Agent-State Visualizer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 9 · Metric: Resume Fidelity

T18 **حارس الانتقال**
Transition Guard
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 9 · Metric: Checkpoint Coverage

T19 **حاسبة ميزانية الأثر**
Effect-Budget Calculator
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 10 · Metric: Reversibility Score

T20 **بوابة التصعيد**
Escalation Gate
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 10 · Metric: Escalation Precision

الملحق ب

فهرس 100 أداة مساعفية · 2/5

Catalog of 100 Moussaif Operational Tools

T21 مُصنّف المهمة الوعائي
Vascular Task Classifier
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 11 · Metric: Mode-Regret

T22 مولّد ملف التشغيل
Runtime-Profile Generator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 11 · Metric: Retyping Latency

T23 موجّه النماذج الفائق
Moussaif HyperRouter
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 12 · Metric: Escalation Yield

T24 حاسبة ندم النموذج
Model-Regret Calculator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 12 · Metric: Fallback Continuity

T25 موجّه HyperSkill
HyperSkill Router
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 13 · Metric: Composition Depth

T26 محلل تعارض المهارات
Skill-Conflict Analyzer
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 13 · Metric: Historical Success Calibration

T27 مجدول الميزانية المعرفية
Cognitive Budget Scheduler
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 14 · Metric: Verification Reserve Use

T28 قاطع الفرع منخفض العائد
Low-Yield Branch Cutter
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 14 · Metric: Deadline Success

T29 متحكم المقاطعات
Agent Interrupt Controller
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 15 · Metric: Duplicate-Action Rate

T30 كاشف تشبع الحلقة
Loop-Saturation Detector
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 15 · Metric: Stop-Regret

T31 موجّه طبقات الذاكرة
Memory-Tier Router
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 16 · Metric: Promotion Error

T32 بوابة ترقية الخبرة
Experience-Promotion Gate
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 16 · Metric: Memory Access Cost

T33 مجزئ صفحات السياق
Context Pager
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 17 · Metric: Compaction Ratio

T34 محكمة عدم الضياع
No-Loss Compaction Court
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 17 · Metric: Page-Fault Utility

T35 مُصرّف KB-as-OS
KB-as-OS Compiler
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 18 · Metric: Policy Coverage

T36 مُنشئ حزمة التنفيذ
Execution-Pack Builder
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 18 · Metric: Context Pack Size

T37 مقطر الخبرة
Episode Distiller
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 19 · Metric: Causal Attribution Accuracy

T38 مولّد الاختبار المضاد
Counterfactual Test Generator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 19 · Metric: Skill Promotion Precision

T39 مخطط احتفاظ الذاكرة
Memory-Retention Planner
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 20 · Metric: Privacy Exposure

T40 ممحاة المشتقات الحساسة
Sensitive-Derivative Eraser
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 20 · Metric: Reconstruction Fidelity

الملحق ب

فهرس 100 أداة مساعفية · 3/5

Catalog of 100 Moussaif Operational Tools

T41 مُصرّف عقد الأداة
Tool-Contract Compiler
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 21 · Metric: False-Success Rate

T42 مفسر أثر الاستدعاء
Call-Effect Interpreter
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 21 · Metric: Rollback Coverage

T43 جسر MCP المساعفي
Moussaif MCP Bridge
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 22 · Metric: Version Break Rate

T44 مصفوفة توافق الأدوات
Tool-Compatibility Matrix
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 22 · Metric: Semantic Consistency

T45 مُثبت الهدف المرئي
Visual Target Anchor
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 23 · Metric: Duplicate Submission

T46 حارس الفعل الخارجي
External-Action Guard
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 23 · Metric: Recovery Steps

T47 قفص التنفيذ
Execution Sandbox
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 24 · Metric: Sandbox Escape Surface

T48 مولّد اختبار الخصائص
Property-Test Generator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 24 · Metric: Deterministic Coverage

T49 موازن ثقة الحواس
Sensor-Confidence Balancer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 25 · Metric: Temporal Alignment

T50 كاشف التعارض متعدد الوسائط
Cross-Modal Conflict Detector
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 25 · Metric: Unsafe Actuation Rate

T51 مُنشئ HyperSkill
HyperSkill Builder
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 26 · Metric: Contract Completeness

T52 ممحص اكتمال المهارة
Skill-Completeness Auditor
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 26 · Metric: Invocation Cost

T53 متحكم Hyperloop
Hyperloop Controller
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 27 · Metric: Loop Cost

T54 مسجل أفضل حالة
Best-State Register
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 27 · Metric: Regression Avoidance

T55 محكمة فجوة القدرة
Capability-Gap Court
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 28 · Metric: Duplicate Skill Rate

T56 مُبلور المهارة
Skill Crystallizer
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 28 · Metric: Boundary-Test Pass

T57 مقعد اختبار HyperSkill
HyperSkill Test Bench
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 29 · Metric: Version Adoption

T58 مسجل نسب المهارة
Skill-Lineage Registry
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 29 · Metric: Contextual Reliability

T59 موجّه الانتشار
Propagation Router
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 30 · Metric: Unauthorized Propagation

T60 مراقب انحراف الحزم
Pack-Drift Monitor
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 30 · Metric: Revocation Latency

الملحق ب

فهرس 100 أداة مساعفية · 4/5

Catalog of 100 Moussaif Operational Tools

T61 حاسبة قيمة التفويض
Delegation-Value Calculator
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 31 · Metric: Coordination Tax

T62 قاطع التكرار الوكيلي
Agent-Duplication Cutter
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 31 · Metric: Duplicate Work

T63 مُركب أدوار الوكلاء
Agent-Role Composer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 32 · Metric: Evidence Duty Fulfillment

T64 مدقق عقد التفويض
Delegation-Contract Auditor
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 32 · Metric: Role Overlap

T65 مُدمج الحالة الموقّتة
Versioned-State Merger
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 33 · Metric: Stale Result Rate

T66 كاشف التعارض السببي
Causal-Conflict Detector
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 33 · Metric: State Convergence

T67 محكمة الوكلاء
Moussaif Agent Court
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 34 · Metric: Decision Calibration

T68 مقياس استقلال الدليل
Evidence-Independence Meter
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 34 · Metric: Debate Token Cost

T69 مُنشئ حزمة القرار
Decision-Packet Builder
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 35 · Metric: Decision Time

T70 سجل التجاوز البشري
Human-Override Ledger
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 35 · Metric: Override Benefit

T71 محكمة الهدم المساعفية
Moussaif Falsification Court
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 36 · Metric: False Acceptance

T72 مولّد شهادة الفشل
Failure-Certificate Generator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 36 · Metric: Domain Qualification

T73 مُحلل متجه الشك
Uncertainty-Vector Analyzer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 37 · Metric: Selective Accuracy

T74 معاير الثقة السياقية
Contextual Confidence Calibrator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 37 · Metric: Uncertainty Resolution Cost

T75 وسيط الصلاحيات
Capability Broker
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 38 · Metric: Taint Propagation

T76 حاجز حقن السياق
Context-Injection Firewall
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 38 · Metric: Blocked Injection Rate

T77 مخطط التراجع
Rollback Planner
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 39 · Metric: Mean Time to Restore

T78 مُعيد الحالة المتسقة
Consistent-State Restorer
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 39 · Metric: Residual Side Effects

T79 مُصرّف سياسة الحوكمة
Governance-Policy Compiler
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 40 · Metric: Auditability

T80 خريطة صاحب الخطر
Risk-Owner Map
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 40 · Metric: Stakeholder Harm Signals

الملحق ب

فهرس 100 أداة مساعفية · 5/5

Catalog of 100 Moussaif Operational Tools

T81 محاسب اقتصاد Token
Token-Economy Accountant
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 41 · Metric: Context Waste

T82 ضاغط حزمة السياق
Context-Pack Compressor
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 41 · Metric: Quality Retention

T83 موازن حصص المزود
Provider-Quota Balancer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 42 · Metric: Fallback Quality

T84 قاطع دائرة النموذج
Model Circuit Breaker
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 42 · Metric: Queue Delay

T85 مُنشئ مفتاح Cache دلالي
Semantic Cache-Key Builder
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 43 · Metric: Stale Hit Rate

T86 محلل المسار الحرج
Critical-Path Analyzer
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 43 · Metric: End-to-End Latency

T87 مرسام الرحلة الوكيلية
Agentic Trace Mapper
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 44 · Metric: Root-Cause Time

T88 منقح السجل الحساس
Sensitive-Log Redactor
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 44 · Metric: Secret Leakage

T89 بوابة نشر الوكيل
Agent Deployment Gate
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 45 · Metric: Change Failure Rate

T90 مولّد Runbook التعافي
Recovery-Runbook Generator
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 45 · Metric: Recovery Time

T91 مُركب مصفوفة Benchmark
Benchmark-Matrix Composer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 46 · Metric: Cost-Normalized Success

T92 كاشف تلوث الاختبار
Benchmark-Contamination Detector
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 46 · Metric: Repeated-Trial Stability

T93 حاسبة MGAI
MGAI Calculator
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 47 · Metric: Safety Floor

T94 مُفسر حدود الثقة
Confidence-Bound Interpreter
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 47 · Metric: Confidence Interval

T95 مختبر Ablation مساعفي
Moussaif Ablation Lab
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 48 · Metric: Interaction Effect

T96 مُشخّص سلسلة الفشل
Failure-Chain Diagnoser
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 48 · Metric: Failure Localization

T97 مصمم دراسة الحالة
Case-Study Designer
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 49 · Metric: User Time Saved

T98 مُراجع فرق العالم
World-Difference Auditor
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 49 · Metric: Error and Harm Rate

T99 سجل الفروض البحثية
Research-Hypothesis Ledger
تحويل أطروحة الفصل إلى قرار Runtime يمكن تتبعه بدل بقائها نصًا إرشاديًا.
الفصل 50 · Metric: Open Failure Reports

T100 خريطة طريق قابلية الهدم
Falsifiable-Roadmap Mapper
قياس الفرق بين النتيجة المتوقعة والمرصودة ومنع الإغلاق الكاذب للمهمة.
الفصل 50 · Metric: Maturity Level

231

الملحق ج

مكتبة 30 Bounded Hyperloops · 1/3

Thirty Governed Hyperloop Patterns

HL1 · حلقة توليد السؤال الأعلى

Higher-Question Loop

سؤال أولي ← بدائل ← Ω(Q) ← اختيار ← إعادة توليد عند نقص الدليل.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL2 · حلقة العقد قبل الفعل

Contract-Before-Action Loop

نية ← قيود ← صلاحية ← شروط توقف ← تثبيت Goal Contract.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL3 · حلقة التخطيط المتحقق

Verified-Planning Loop

خطة ← Preconditions ← محاكاة أثر ← نقد ← خطة موقعة.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL4 · حلقة ReAct المحكومة

Governed ReAct Loop

Reason ← Guarded Action ← Observation ← State Update ← Verification.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL5 · حلقة محكمة الهدم

Falsification-Court Loop

ادعاء ← أقوى ناقض ← اختبار مستقل ← تقييد أو رفض أو بقاء.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL6 · حلقة ترقية الذاكرة

Memory-Promotion Loop

Episode ← تقطير ← Counterfactual ← إعادة إنتاج ← Procedure.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL7 · حلقة الضغط القابل للفك

Reversible-Compaction Loop

سياق ← Claims/Constraints ← موجز مؤشّر ← اختبار استرجاع.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL8 · حلقة توجيه النموذج

Model-Routing Loop

خطوة ← تعقيد/خطر ← Tier ← نتيجة ← Regret ← تحديث Router.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL9 · حلقة توجيه المهارة

Skill-Routing Loop

Trigger ← Candidates ← Compatibility ← Composition ← Tests.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL10 · حلقة اقتصاد الحوسبة

Compute-Economy Loop

ميزانية ← عائد هامشي ← حجز تحقق ← قطع فرع ← إعادة توزيع.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

قاعدة مشتركة كل Hyperloop تحفظ أفضل حالة معروفة، وتصدر سبب التوقف، ولا ترفع صلاحيتها تلقائيًا أثناء الفشل.

الملحق ج

مكتبة 30 Bounded Hyperloops · 2/3

Thirty Governed Hyperloop Patterns

HL11 · حلقة التصعيد البشري

Human-Escalation Loop

خطر ← قيمة تدخل ← Decision Packet ← Override ← تعلم سياسة.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL12 · حلقة الاسترداد

Recovery Loop

فشل ← تجميد أثر ← تشخيص ← Rollback/Compensation ← تحقق حالة.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL13 · حلقة أمن الأداة

Tool-Security Loop

Revoke. ← Postcondition ← Call ← Capability ← Taint ← Schema

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL14 · حلقة المتصفح المفسرة

Explainable-Browser Loop

Verify UI. ← Click/Type ← Guard ← Anchor ← Observe

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL15 · حلقة الدمج متعدد الوكلاء

Multi-Agent Merge Loop

Versioned Merge. ← Judge ← Conflicts ← Artifacts ← Snapshots

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL16 · حلقة توليد الدور

Role-Synthesis Loop

Evidence Packet. ← Agent ← Contract ← Gap ← Task Graph

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL17 · حلقة تشبع البحث

Search-Saturation Loop

Stop or Continue. ← Cost/Risk ← Evidence Gain ← فرضيات

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL18 · حلقة إصدار HyperSkill

HyperSkill Release Loop

Promote/Rollback. ← Canary ← Regression ← Sandbox ← Candidate

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL19 · حلقة الانتشار التناسلي

Propagation Loop

Drift Monitor. ← Pack ← Compatibility ← Project Scope ← Lineage

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL20 · حلقة معايرة الثقة

Confidence-Calibration Loop

Policy Update. ← Calibration Bin ← توقع ← نتيجة

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

قاعدة مشتركة كل Hyperloop تحفظ أفضل حالة معروفة، وتصدر سبب التوقف، ولا ترفع صلاحيتها تلقائيًا أثناء الفشل.

الملحق ج

مكتبة 30 Bounded Hyperloops · 3/3

Thirty Governed Hyperloop Patterns

HL21 · حلقة مراقبة الانجراف

Drift-Monitoring Loop

Retype/Retrain/Escalate. ← Shift Test ← Live Trace ← Baseline

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL22 · حلقة صحة KB-as-OS

KB Health Loop

Deprecation. ← Validity ← Contradictions ← Provenance ← Nodes

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL23 · حلقة النقل بين المجالات

Cross-Domain Transfer Loop

Negative Transfer Test. ← Adapt ← New Domain ← Invariant

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL24 · حلقة بناء Benchmark

Benchmark-Design Loop

Contamination Audit. ← Metric ← Holdout ← Task ← Capability Claim

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL25 · حلقة Ablation

Ablation Loop

Causal Delta. ← Equal Budget ← Remove One Service ← Full System

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL26 · حلقة الأثر العالمي

World-Effect Loop

Externalities. ← Verified Delta ← Observation ← Action ← Desired State

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL27 · حلقة احترام RPM

RPM-Aware Loop

Recovery. ← Fallback ← Backoff ← Priority ← Queue ← Quota

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL28 · حلقة الاستشعار المتعدد

Multimodal-Fusion Loop

Act/Wait. ← Conflict ← Confidence ← Time Align ← Sensors

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL29 · حلقة الدستور التنفيذي

Runtime-Constitution Loop

Regression. ← Action Audit ← Preflight ← Compile ← Rules

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

HL30 · حلقة الذكاء الفتّاح

Open-Intelligence Loop

Governed ← Falsification ← New Tool/Skill ← Higher Question ← Gap Growth.

حد الحلقة: Iteration + Cost + Risk + Evidence Gain.

قاعدة مشتركة كل Hyperloop تحفظ أفضل حالة معروفة، وتصدر سبب التوقف، ولا ترفع صلاحيتها تلقائيًا أثناء الفشل.

الملحق د

Rubric النشر ذي 26 نقطة · 1/3

Moussaif 26-Point Deployment Rubric

1. تحديد Owner وغاية العمل والأثر المرغوب.
2. تحويل الطلب إلى Goal Contract قابل للاختبار.
3. تصنيف المهمة ونوع الحداثة وحساسية البيانات.
4. تحميل الدستور التنفيذي وإجراء Constitution Preflight.
5. إنشاء World State وBelief State منفصلين.
6. اختيار أقل Context Pack يحقق التغطية.
7. استرجاع HyperSkills مع شروطها وإصداراتها.
8. حساب Model Route على الجودة والكلفة والخطر.
9. تخصيص Budget للتحقق قبل الفعل.

بوابة الصفحة

PASS دليل أو Artifact لكل نقطة.

FAIL نقطة حساسة غير مثبتة توقف الترقية.

لا يسمح المتوسط بإخفاء فشل الصلاحيات أو الاسترداد أو أثر خارجي غير مصرح. تحفظ النتيجة مع Runtime Version وModel Route وTool Versions.

Deploy iff Coverage = 26/26 or an Explicit, Owned, Time-Bounded Exception Exists

الملحق د

Rubric النشر ذي 26 نقطة · 2/3

Moussaif 26-Point Deployment Rubric

10. بناء Plan Graph وتحديد Critical Path.
11. تعليم الأفعال الخارجية وغير القابلة للعكس.
12. تحديد الصلاحيات الأدنى ومدة صلاحيتها.
13. إنشاء Rollback أو Compensating Action.
14. تفويض Subagents فقط عند قيمة موجبة.
15. تثبيت عقود المخرجات للوكلاء الفرعيين.
16. تشغيل الأدوات عبر Schemas وتحقق مسبق.
17. تحديث State بعد Observation لا بعد التخمين.
18. تشغيل Falsification Court على الادعاءات الحساسة.

بوابة الصفحة

PASS دليل أو Artifact لكل نقطة.

FAIL نقطة حساسة غير مثبتة توقف الترقية.

لا يسمح المتوسط بإخفاء فشل الصلاحيات أو الاسترداد أو أثر خارجي غير مصرح. تحفظ النتيجة مع Runtime Version وModel Route وTool Versions.

Deploy iff Coverage = 26/26 or an Explicit, Owned, Time-Bounded Exception Exists

الملحق د

Rubric النشر ذي 26 نقطة · 3/3

Moussaif 26-Point Deployment Rubric

19. معايرة عدم اليقين وتحديد ما يحله.
20. طلب Human Override عند عبور عتبة الخطر.
21. التحقق من Postconditions والأثر العالمي.
22. حفظ Trace من الهدف إلى الدليل.
23. تقطير Episode من دون تعميم متسرع.
24. ترقية Skill فقط بعد إعادة إنتاج واختبار حدودي.
25. قياس الجودة والكلفة والزمن والمخاطرة معًا.
26. إغلاق المهمة بشهادة نجاح أو فشل وحدود صلاحية واضحة.

بوابة الصفحة

PASS دليل أو Artifact لكل نقطة.

FAIL نقطة حساسة غير مثبتة توقف الترقية.

لا يسمح المتوسط بإخفاء فشل الصلاحيات أو الاسترداد أو أثر خارجي غير مصرح. تحفظ النتيجة مع Runtime Version وModel Route وTool Versions.

Deploy iff Coverage = 26/26 or an Explicit, Owned, Time-Bounded Exception Exists

الملحق هـ

بروتوكولات Benchmark · 1/3

Benchmark and Evaluation Protocols

GAIA

General AI Assistant tasks

القدرة: Multimodality + Tools + Web + Reasoning

المقاييس: Exact answer، Tool trace، Cost، Attempts

ضبط الصدق: اختبر Holdout ولا تسمح بتسرب الإجابة في KB.

المرجع: R19

AgentBench

Interactive agent environments

القدرة: Long-horizon decision and instruction following

المقاييس: Success، Steps، Invalid Actions، Recovery

ضبط الصدق: سجل نوع الفشل لا النتيجة النهائية فقط.

المرجع: R18

قالب التقرير الإلزامي

نسخة البيئة	Dataset/Commit/Date وصلاحيات الأدوات.
ميزانية التشغيل	Models، Calls، Tokens، Time، Money، Retries.
المخرجات	Outcome Score + Trajectory Score + Failure Chain.
السلامة	Unauthorized Effects، Secrets، Rollback، Human Override.
الثقة	Seeds، Variance، Confidence Interval، Negative Results.

Fair Comparison Requires Equal Authority, Equal Budget, and Equal Information

الملحق هـ

بروتوكولات Benchmark · 2/3

Benchmark and Evaluation Protocols

WebArena

Realistic web tasks

القدرة: Navigation, grounding, functional completion

المقاييس: End-to-end success، Duplicate Effects، Latency

ضبط الصدق: استخدم بيئة ذاتية الاستضافة وتحققًا وظيفيًا.

المرجع: R17

OSWorld

Real computer environments

القدرة: GUI grounding across desktop applications

المقاييس: Execution success، Grounding، Time، Safety

ضبط الصدق: اعزل VM واختبر Rollback لكل فعل حساس.

المرجع: R20

قالب التقرير الإلزامي

نسخة البيئة	Dataset/Commit/Date وصلاحيات الأدوات.
ميزانية التشغيل	Models، Calls، Tokens، Time، Money، Retries.
المخرجات	Outcome Score + Trajectory Score + Failure Chain.
السلامة	Unauthorized Effects، Secrets، Rollback، Human Override.
الثقة	Seeds، Variance، Confidence Interval، Negative Results.

Fair Comparison Requires Equal Authority, Equal Budget, and Equal Information

الملحق هـ

بروتوكولات Benchmark · 3/3

Benchmark and Evaluation Protocols

SWE-bench

Repository issue resolution

القدرة: Code understanding, editing, and tests

المقاييس: Resolved، Test delta، Patch size، Cost

ضبط الصدق: ثبّت نسخة المستودع ومنع الوصول إلى حلول الاختبار.

المرجع: R21

Moussaif Cross-Domain Suite

Research + Code + Business + Documents

القدرة: Transfer, governance, recovery, and Arabic RTL

المقاييس: MGAI vector، World Effect، Human time، Risk

ضبط الصدق: مجموعة جديدة مقترحة؛ يجب نشر المهام ومراجعتها خارجيًا.

المرجع: R25

قالب التقرير الإلزامي

نسخة البيئة	Dataset/Commit/Date وصلاحيات الأدوات.
ميزانية التشغيل	Models، Calls، Tokens، Time، Money، Retries.
المخرجات	Outcome Score + Trajectory Score + Failure Chain.
السلامة	Unauthorized Effects، Secrets، Rollback، Human Override.
الثقة	Seeds، Variance، Confidence Interval، Negative Results.

Fair Comparison Requires Equal Authority, Equal Budget, and Equal Information

الملحق و

Pseudocode النواة التنفيذية

Reference Kernel Pseudocode

```
contract = compile_intent(user_request, authority, constraints)
question = govern_question(contract, q_mcd_profile)
state = load_world_belief_desired_state(project_registry)
runtime = type_task(question, effect, freshness, horizon)
pack = compile_kb_os(runtime, least_context = true)
route = hyperroute(models, skills, tools, budget, risk)
plan = build_plan(route, guards, rollback, stop_rules)
while not done and evidence_gain > marginal_cost:
action = select_guarded_action(plan, state)
capability = authorize_minimum(action)
result = execute_tool_contract(action, capability)
observation = verify_postcondition(result)
state = transition(state, observation)
verdict = falsification_court(state, plan)
plan = update_or_recover(verdict, best_checkpoint)
artifact = commit_only_if_verified(state, contract)
learn = distill_episode_without_unverified_generalization()
```

Invariants

- Constitution precedes model output
- Observation precedes state commit
- Verification reserve cannot be spent early
- Best checkpoint survives every loop

Terminal states

- Verified Commit
- Bounded Partial Result
- Rolled Back
- Escalated with Decision Packet

Implementation note يمكن تحقيق النواة كـState Machine وخدمات مستقلة، أو داخل Orchestrator واحد في الإصدار الأول، بشرط بقاء الحدود قابلة للاختبار.

الملحق و

Schema مرجعي لـHyperSkill

Reference HyperSkill Schema

```
name: repository-domain-migration
version: 1.2.0
purpose: replace a canonical domain safely across active repositories
triggers: replace URL | migrate domain | update canonical link
requires: project_registry, shell_read, scoped_write, tests
preconditions: old_value, new_value, active_repo_scope
procedure: inventory → patch → scan → test → deploy_if_authorized
verification: old_count=0, tests=pass, generated_files=inspected
failure_modes: binary, archive, encoded_url, generated_artifact
rollback: restore versioned files and deployment checkpoint
metrics: coverage, regression_delta, execution_cost
lineage: moussaif-web-domain-migration · version 1.2.0
```

قواعد البناء

- اسم يعبر عن قدرة لا دور تمثيلي
- Purpose يذكر الأثر
- Triggers محددة
- Requires بأقل صلاحية
- Verification مستقل عن المولد
- Failure Modes من Episodes حقيقية
- Cost Profile لكل مزود
- Lineage وVersion
- Rollback مجرب
- Promotion بعد Canary

Reusable Skill = Stable Interface + Bounded Procedure + Independent Verification

الملحق و

Router والصلاحيات وآلة الحالة

Routing, Permissions, and State Machine

HyperRouting Score

Rel + Utility + Compat + Evidence + History - Cost - Risk - Context

يعاد الحساب عند فشل أداة أو تغير Deadline أو ورود رسالة تقلب نوع المهمة.

Effect Authorization

Permit = Need × Scope × FreshAuth × TaintSafe

إذا كان أي عامل صفرًا يمنع الفعل أو يصعده؛ لا يعوض عامل عاملًا.

آلة الحالة القياسية

INTAKE
نية خام وسلطة أولية.

CONTRACTED
هدف وقيود وتوقف.

PLANNED
مسار وحراس واسترداد.

ACTING
Capability قصيرة العمر.

VERIFYING
Postconditions ومحكمة هدم.

COMMITTED
أثر موثق وArtifact.

PAUSED
Checkpoint ومقاطعة.

ROLLED BACK
حالة متسقة مستعادة.

ESCALATED
Decision Packet للإنسان.

قاعدة الانتقال لا يوجد انتقال مباشر من PLANNED إلى COMMITTED. الفعل والملاحظة والتحقق خدمات مستقلة حتى لو نفذها النموذج نفسه في إصدار أولي.

الملحق ز

المعجم العربي-الإنجليزي · 1/3

Arabic-English Technical Glossary

Agent وكيل
كيان يدير هدفًا وحالة وأفعالًا ضمن حدود.

Agentic وكيلي/ذو وكالة
متعلق بالقدرة على التخطيط والفعل والملاحظة.

AGI-level ذو مستوى مستهدف للعمومية
معيار بحثي متعدد الأبعاد، لا ادعاء بتحقق AGI.

Runtime زمن التشغيل
البيئة الفعلية التي تنفذ فيها القواعد والأدوات.

Kernel النواة
طبقة الحاكمية والحالة والجدولة والصلاحيات.

HyperRouting التوجيه الفائق
اختيار النموذج والمهارة والمسار والميزانية.

HyperSkill مهارة فائقة
وحدة قدرة ذات عقد واختبارات وإصدار.

Hyperloop حلقة فائقة
دورة تعلم أو تنفيذ محدودة بالدليل والكلفة والخطر.

KB-as-OS قاعدة المعرفة كنظام تشغيل
معرفة تفعّل سياسات ومهارات واختبارات في Runtime.

Context Pack حزمة سياق
أقل مجموعة معرفة وقيود لازمة للمهمة.

Working Memory ذاكرة عاملة
الحالة الساخنة اللازمة للخطوة الحالية.

Episodic Memory ذاكرة حدثية
سجل محاولات ونتائج مرتبطة بسياقها.

Semantic Memory ذاكرة دلالية
ادعاءات ومفاهيم مستقرة نسبيًا.

Procedural Memory ذاكرة إجرائية
مهارات وخطوات تشغيل موثقة.

Provenance سلسلة المصدر
أصل الادعاء وتحولاته وإصداره.

Claim ادعاء
وحدة معرفة قابلة للدعم أو النقض.

Observation ملاحظة
بيان ورد من أداة أو بيئة مع مصدر وزمن.

Inference استنتاج
نتيجة مشتقة من ملاحظات وقواعد.

World State حالة العالم
ما رُصد بوصفه قائمًا الآن.

Belief State حالة الاعتقاد
تقدير النظام لما يرجح أنه قائم.

الملحق ز

المعجم العربي-الإنجليزي · 2/3
Arabic-English Technical Glossary

Desired State الحالة المطلوبة
النتيجة التي يحددها Goal Contract.

Guard حارس انتقال
شرط يمنع فعلًا أو انتقالًا غير صالح.

Postcondition شرط لاحق
حقيقة يجب رصدها بعد نجاح الفعل.

Idempotency عدم تضاعف الأثر
إعادة الاستدعاء لا تنشئ أثرًا إضافيًا.

Rollback تراجع
إعادة الحالة إلى نسخة موثوقة.

Compensation تعويض
فعل مضاد يحد أثر فعل لا يمكن عكسه حرفيًا.

Falsification تفنيد/هدم
محاولة منظمة لإيجاد ناقض للفرضية.

Calibration معايرة
مطابقة الثقة المعلنة مع التكرار الفعلي للصواب.

Uncertainty Vector متجه عدم اليقين
فصل الشك المعرفي والرصدّي والأداتي والزمني والسببي.

Capability Token رمز قدرة
إذن محدود النطاق والزمن لفعل معين.

Taint Tracking تتبع تلوث البيانات
وسم البيانات غير الموثوقة ومنع تحولها إلى أوامر.

MCP بروتوكول سياق النموذج
معيار لربط التطبيقات بالأدوات والموارد.

Schema مخطط
تعريف منظم للمدخلات والمخرجات والقيود.

Adapter مهايئ
طبقة تحول واجهة أو دلالة إلى عقد موحد.

Benchmark محك اختباري
مجموعة مهام ومقاييس لمقارنة أنظمة.

Ablation اختبار إزالة
حذف مكوّن لقياس مساهمته السببية.

Canary نشر كناري
تجربة إصدار على نسبة صغيرة قبل التعميم.

SLO هدف مستوى الخدمة
حد معلن للجودة أو التوفر أو الزمن.

Critical Path المسار الحرج
أطول سلسلة تبعيات تحد زمن الإكمال.

Circuit Breaker قاطع دائرة
إيقاف مؤقت لمورد يفشل لتجنب انهيار متكرر.

الملحق ز

المعجم العربي-الإنجليزي · 3/3
Arabic-English Technical Glossary

Rate Limit حد المعدل
سقف الاستدعاءات أو Tokens في زمن معين.

Token Economy اقتصاد الرموز
تخصيص Tokens وفق أثر موثق لا طول النص.

Semantic Cache تخزين دلالي مؤقت
إعادة استخدام نتيجة متوافقة في المعنى والإصدار.

Trace أثر تتبع
سلسلة مترابطة من الهدف إلى الفعل والدليل.

Checkpoint نقطة حفظ
حالة تسمح باستئناف آمن بعد توقف.

Snapshot لقطة
نسخة حالة محددة الإصدار والزمن.

Subagent وكيل فرعي
عامل مفوض بمهمة مستقلة وعقد محدود.

Orchestrator منسق
يدير التبعيات والتفويض والدمج والحكم.

Coordination Tax ضريبة التنسيق
كلفة الرسائل والدمج والتعارض بين الوكلاء.

Decision Packet حزمة قرار
ملخص بدائل وأدلة وآثار لتدخل الإنسان.

Human-in-the-Loop الإنسان داخل الحلقة
موافقة أو حكم في خطوة تنفيذية حساسة.

Human-over-the-Loop الإنسان فوق الحلقة
وضع السياسة والمراقبة والمساءلة.

Transfer نقل القدرة
تطبيق ثابت معرفي في مجال جديد مع تكييف القيود.

Negative Transfer نقل سلبي
انخفاض الأداء بسبب تشابه سطحي مضلل.

Generality Surface سطح العمومية
متجه النقل والتكيف والتحقق والكفاءة والسلامة.

MGAI مؤشر الوكالة العامة المساعفي
تركيب مشروط ببوابات سلامة لمقاييس الوكالة.

MCD الفرق المعرفي المساعفي
تجميع الفرق والتكامل والزمن والوكالة والغاية والسياق والضوضاء.

Q-MCD MCD المحكوم بالسؤال
ضرب MCD في جودة السؤال $\Omega(Q)$.

Open Intelligence الذكاء الفتّاح
أفق يولد أسئلة وقدرات جديدة تحت الحاكمية.

M-AGIOS نظام التشغيل المساعفي
الإطار النظري والتنفيذي المقترح في هذا الكتاب.

246

المراجع

المصادر الأولية والتأسيسية · 1/3

Primary and Foundational Bibliography

Alan M. Turing. Computing Machinery and Intelligence. Mind 59(236), 1950.
https://doi.org/10.1093/mind/LIX.236.433 R01

Allen Newell and Herbert A. Simon. Human Problem Solving. Prentice-Hall, 1972.
https://archive.org/details/humanproblemsolv0000newe R02

Michael Bratman. Intention, Plans, and Practical Reason. Harvard University Press, 1987.
https://www.hup.harvard.edu/books/9780674458185 R03

Richard S. Sutton and Andrew G. Barto. Reinforcement Learning: An Introduction. MIT Press, 2018.
http://incompleteideas.net/book/the-book-2nd.html R04

Stuart Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. Pearson, 4th edition, 2020.
https://aima.cs.berkeley.edu/ R05

Ashish Vaswani et al.. Attention Is All You Need. NeurIPS, 2017.
https://arxiv.org/abs/1706.03762 R06

Jason Wei et al.. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. NeurIPS, 2022.
https://arxiv.org/abs/2201.11903 R07

Xuezhi Wang et al.. Self-Consistency Improves Chain of Thought Reasoning. ICLR, 2023.
https://arxiv.org/abs/2203.11171 R08

Shunyu Yao et al.. ReAct: Synergizing Reasoning and Acting in Language Models. ICLR, 2023.
https://arxiv.org/abs/2210.03629 R09

Timo Schick et al.. Toolformer: Language Models Can Teach Themselves to Use Tools. NeurIPS, 2023.
https://arxiv.org/abs/2302.04761 R10

منهج الاستعمال المراجع تسند السياق العلمي والأنماط السابقة. الأسماء والمعادلات المساعفية الجديدة تبقى مقترحات هذا الكتاب وتحتاج اختبارات مستقلة.

247

المراجع

المصادر الأولية والتأسيسية · 2/3

Primary and Foundational Bibliography

Noah Shinn et al.. Reflexion: Language Agents with Verbal Reinforcement Learning. NeurIPS, 2023.
https://arxiv.org/abs/2303.11366 R11

Shunyu Yao et al.. Tree of Thoughts: Deliberate Problem Solving with Large Language Models. NeurIPS, 2023.
https://arxiv.org/abs/2305.10601 R12

Joon Sung Park et al.. Generative Agents: Interactive Simulacra of Human Behavior. UIST, 2023.
https://arxiv.org/abs/2304.03442 R13

Guanzhi Wang et al.. Voyager: An Open-Ended Embodied Agent with Large Language Models. 2023.
https://arxiv.org/abs/2305.16291 R14

Charles Packer et al.. MemGPT: Towards LLMs as Operating Systems. 2023.
https://arxiv.org/abs/2310.08560 R15

Qingyun Wu et al.. AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation. 2023.
https://arxiv.org/abs/2308.08155 R16

Shuyan Zhou et al.. WebArena: A Realistic Web Environment for Building Autonomous Agents. ICLR, 2024.
https://arxiv.org/abs/2307.13854 R17

Xiao Liu et al.. AgentBench: Evaluating LLMs as Agents. ICLR, 2024.
https://arxiv.org/abs/2308.03688 R18

Grégoire Mialon et al.. GAIA: A Benchmark for General AI Assistants. ICLR, 2024.
https://arxiv.org/abs/2311.12983 R19

Tianbao Xie et al.. OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments. NeurIPS, 2024.
https://arxiv.org/abs/2404.07972 R20

منهج الاستعمال المراجع تسند السياق العلمي والأنماط السابقة. الأسماء والمعادلات المساعفية الجديدة تبقى مقترحات هذا الكتاب وتحتاج اختبارات مستقلة.

المراجع

المصادر الأولية والتأسيسية · 3/3

Primary and Foundational Bibliography

Carlos E. Jimenez et al.. SWE-bench: Can Language Models Resolve Real-World GitHub Issues?. ICLR, 2024. https://arxiv.org/abs/2310.06770 R21

Patrick Lewis et al.. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS, 2020. https://arxiv.org/abs/2005.11401 R22

Long Ouyang et al.. Training Language Models to Follow Instructions with Human Feedback. NeurIPS, 2022. https://arxiv.org/abs/2203.02155 R23

Yuntao Bai et al.. Constitutional AI: Harmlessness from AI Feedback. 2022. https://arxiv.org/abs/2212.08073 R24

NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1, 2023. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf R25

Model Context Protocol. MCP Specification 2025-11-25. Official specification. https://modelcontextprotocol.io/specification/2025-11-25 R26

Wujiang Xu et al.. A-Mem: Agentic Memory for LLM Agents. 2025. https://arxiv.org/abs/2502.12110 R27

Zhao Yang et al.. Language Agent Tree Search Unifies Reasoning, Acting, and Planning. 2023. https://arxiv.org/abs/2310.04406 R28

Sirui Hong et al.. MetaGPT: Meta Programming for Multi-Agent Collaborative Framework. ICLR, 2024. https://arxiv.org/abs/2308.00352 R29

Chen Qian et al.. ChatDev: Communicative Agents for Software Development. ACL, 2024. https://arxiv.org/abs/2307.07924 R30

منهج الاستعمال المراجع تسند السياق العلمي والأنماط السابقة. الأسماء والمعادلات المساعفية الجديدة تبقى مقترحات هذا الكتاب وتحتاج اختبارات مستقلة.

الملحق ح

فهرس المعادلات والمؤشرات

Equation and Indicator Index

MCD = λ1D + λ2I + λ3T + λ4A + λ5P + λ6C - λ7N

ToolSuccess = TransportOK × PostconditionOK × EvidenceBound

Q-MCD = Ω(Q) × MCD

Delegate iff ParallelGain + SpecialtyGain > CoordinationCost + MergeRisk

OS_valid = State × Schedule × Isolation × Recovery × Observe

R_effect = Uncertainty × Impact × (1 - Reversibility) × Sensitivity

G_surface = GeoMean(Transfer, Adaptation, Verification, Efficiency, Safety)

Keep(m) iff FutureUtility > Storage + PrivacyRisk + DriftRisk

E_claim = SourceQuality × Independence × Recency × Falsifiability

EffectDensity = VerifiedWorldEffect / (Tokens + Cost + Latency + Risk)

Route* = argmax(Quality - Cost - Risk - Latency)

MGAI = SafetyGate × GeoMean(T, A, V, E, R, G, H)

Score(s) = Relevance + Utility + Compatibility + Evidence + History - Costs

Impact = Verified(DesiredState - BaselineState) - Externalities

Stop iff E(ΔEvidence_next) ≤ Cost_next + Risk_next

تنبيه منهجي هذه صيغ Operational للتصميم والمقارنة. لا تصبح Laws تجريبية إلا بعد تعريف المقاييس ومعايرة الأوزان وتكرار النتائج عبر بيئات مستقلة.

الخاتمة

ميثاق مساعف للوكالة المحكومة

Moussaif Manifesto for Governed Agency

لا تُوسّع القدرة قبل أن تُوسّع الحاكمية

Do not scale agency faster than evidence, authority, and recovery.

1. السؤال الحاكم يسبق الجواب البارع.
2. العالم المرصود يسبق الاعتقاد المريح.
3. عقد الأداة يسبق استدعاءها.
4. الدليل المستقل يعلو على صدى الوكلاء.
5. الذاكرة تُستقلب ولا تتراكم بلا نهاية.
6. الحلقة تُحد بزيادة الدليل لا بالرغبة في الاستمرار.
7. التفويض امتياز تحكمه أقل صلاحية.
8. الفعل القابل للعكس هو الافتراض الأول.
9. النجاح أثر موثق لا نص نهائي.
10. المؤشر لا يعوض فشل السلامة.
11. النتيجة السلبية جزء من التقدم.
12. الذكاء الفتّاح نمو محكوم لا ادعاء مطلق.

الخلاصة المؤسِّسة M-AGIOS ليس عقلًا سريًا داخل النموذج؛ إنه دستور وذاكرة وجدولة وعقود وأدلة وحلقات واسترداد حول نماذج قابلة للتبديل. قوته الحقيقية لا تقاس بعدد الملفات أو أسماء الوكلاء، بل بقدرته على حفظ الغاية وإحداث فرق صالح وإثباته والرجوع عنه حين يلزم.

إنجاز وتركيب زهير مساعف

Zouhair Moussaif · Moussaif AGI Research Program

Founding Edition 1.0 · 13 July 2026 · 250 pages