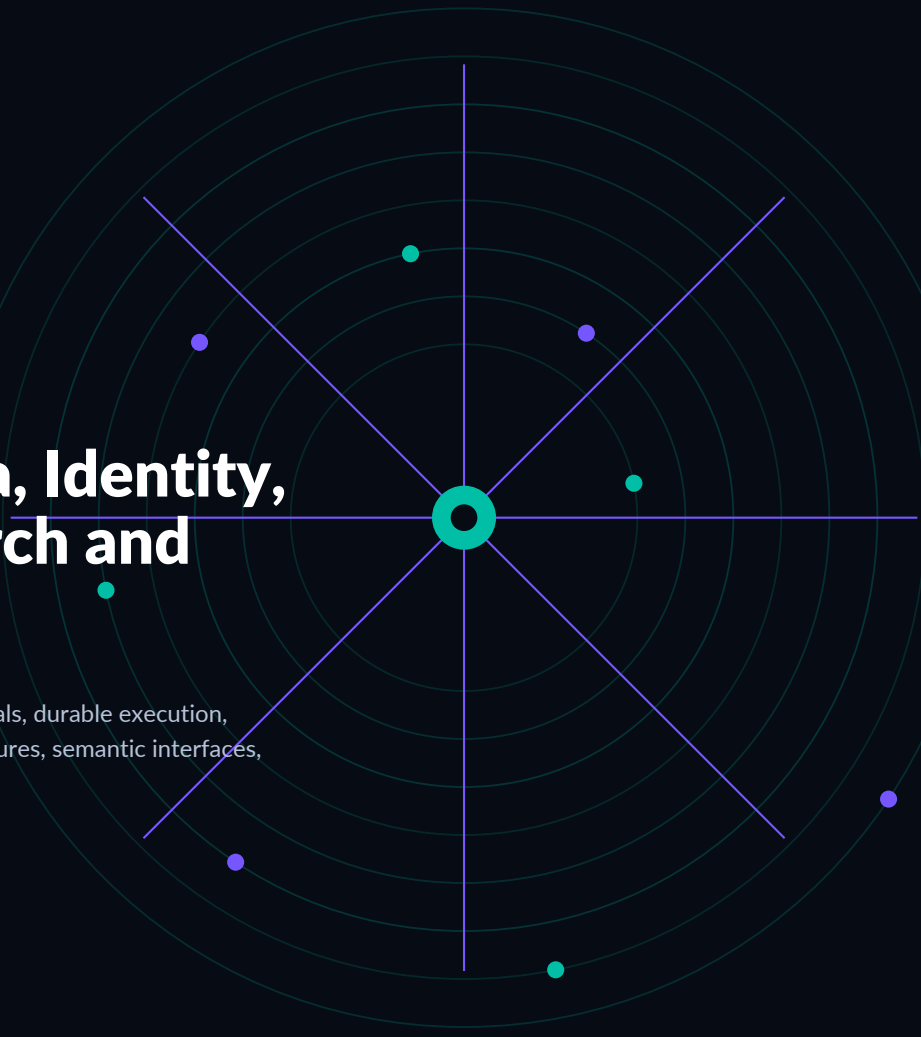
M MYTHOS SYSTEMS

ENGINEERING STANDARDS LIBRARY / RESEARCH ANNEX

AI, Knowledge, Data, Identity, and Learning Research and Adoption Annex

Source-controlled research on agent memory, credentials, durable execution, observability, model routing, advanced model architectures, semantic interfaces, and local-first state.



PERMANENT DOCUMENT ID
MYTHOS-RES-ANNEX-AIKDIL-0001

PUBLICATION
Research Annex

STATUS
Candidate Research Annex

Contents

Research governance

Research strength and adoption status are separate decisions. Documentation, demonstrations, and vendor claims remain below reproduced laboratory evidence. Every adoption posture has a review date, disconfirming condition, and rollback requirement.

Research adoption register

RESEARCH NOTE	ADOPTION POSTURE	KEY LIMITATION
RES-002: Agent Memory Poisoning (Vectors and Mitigations)	Pilot with guardrails	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-003: Credential Brokering for Agents (Zero-Knowledge Secrets)	Pilot with guardrails	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-006: Browser-Local WebLLM (Architecture, Constraints, and WebGPU bindings)	Pilot with guardrails	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-007: Durable Multi-Agent Execution (Temporal/Restate patterns)	Adopt as foundation	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-009: Evidence Bundles (Cryptographic Provenance for Agent Actions)	Adopt as foundation	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-010: OpenTelemetry for Agents (Semantic mapping for cognitive traces)	Pilot with guardrails	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-012: AI Model Routing & Task Placement (Local vs. Hosted heuristics)	Adopt as foundation	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-013: Mixture of Experts (MoE) Routing Dynamics	Benchmark and monitor	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-014: State Space Models (SSMs/Mamba) & Infinite Context	Benchmark and monitor	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-015: Liquid Neural Networks (LNNs)	Benchmark and monitor	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.
RES-016: Neuro-Symbolic AI Integration	Benchmark and monitor	Documentation and research evidence must be reproduced against the intended workload before consequential adoption.

Evidence grades

GRADE	EVIDENCE	ADOPTION LIMIT
A	Reproduced in the intended environment with controlled tests, retained artifacts, and independent review.	May support production adoption within the tested profile and stated limitations.
B	Reproduced prototype or accepted enterprise proof of concept with partial operational evidence.	Conditional adoption with monitored rollout and explicit residual risk.
C	Official specifications, primary documentation, credible research, or vendor evidence without local reproduction.	Pilot, benchmark, or architecture planning only for high-consequence use.
D	Secondary analysis, demonstrations, preliminary studies, or incomplete implementation claims.	Research and controlled experimentation only.
E	Unverified assertion, anecdote, marketing claim, or unresolved conflicting evidence.	Do not use as adoption authority.

Freshness and expiration model

SOURCE VOLATILITY	DEFAULT REVIEW	EXAMPLES
Continuous	30 days	Model endpoints, managed-agent services, security advisories, rolling catalogs, and active preview APIs.
Dynamic	90 days	Agent protocols, OpenTelemetry GenAI conventions, browser AI APIs, fast-moving framework behavior.
Periodic	180 days	Product releases, implementation guidance, benchmark suites, and ecosystem standards.
Stable	365 days	Candidate Standards and mature protocols unless supersession or errata occurs.
Event-triggered	Immediate	Material vulnerability, model or provider change, policy change, failed experiment, or contradictory primary evidence.

Adoption decisions and guardrails

DECISION	WHEN APPROPRIATE	REQUIRED GUARDRAIL
Adopt as foundation	Mature pattern with strong evidence and broad architectural value.	Version control, ownership, operational SLOs, and continuous revalidation.
Conditional adoption	Value is clear but workload, integration, or risk remains context-dependent.	Restricted profile, acceptance criteria, monitoring, rollback, and expiration.
Pilot	Evidence is promising but insufficient for consequential production use.	Isolated environment, synthetic or non-sensitive data, bounded cost, and documented learning goals.
Benchmark and monitor	Technology may matter but current fit or maturity is uncertain.	Repeatable workload tests, source watch, and explicit trigger for reconsideration.
Research only	Scientific or technical feasibility remains unsettled.	No production authority, protected data, or critical dependency.
Do not adopt	Mandatory gate fails or risk exceeds benefit.	Record rationale, disconfirming conditions, and next review trigger.

Explicit research limitations

LIMITATION CLASS	QUESTION TO RECORD
Reproducibility	Was the claimed result reproduced on the intended hardware, data, model, framework, and workload?
External validity	Does the source population, benchmark, or environment represent the intended deployment?
Security	Were poisoning, prompt injection, extension compromise, secret exposure, and privilege escalation tested?
Privacy and sovereignty	Were embeddings, telemetry, memory, model providers, and data movement included in the analysis?
Interoperability	Are protocol, framework, model, storage, and observability claims validated across versions and vendors?
Operational maturity	Are failure handling, upgrades, rollback, support, capacity, and incident response demonstrated?
Economic validity	Are total compute, storage, integration, staffing, egress, and replacement costs measured?
Temporal validity	What source, release, threat, or market change invalidates the conclusion?

Priority validation experiments

WORKSTREAM	MINIMUM EXPERIMENT
Memory poisoning	Inject conflicting, malicious, stale, and cross-tenant memories; verify quarantine, provenance, correction, and deletion.
Credential brokerage	Demonstrate that secrets never enter model-visible context; test audience, expiration, revocation, and confused-deputy resistance.
Durable execution	Kill processes and dependencies during plans, approvals, tools, and commits; verify replay and exactly-once side-effect controls.
Agent observability	Trace model, retrieval, memory, tool, policy, human, cost, and outcome across asynchronous boundaries.
Model routing	Compare quality, safety, latency, privacy, and cost across task profiles and failure conditions.
Advanced architectures	Benchmark MoE, state-space, liquid, and neuro-symbolic approaches against strong conventional baselines.
Semantic interfaces	Validate typed rendering, untrusted content isolation, accessibility, version negotiation, and human override.
Local-first state	Test offline mutation, conflict, merge, migration, encryption, deletion, and recovery across nodes.

Revalidation triggers

- A relevant model, provider, framework, protocol, source, or product changes materially.
- A vulnerability, data incident, policy decision, or failed experiment contradicts an assumption.
- A benchmark becomes unrepresentative of the intended workload or operational environment.
- A new primary source changes the scientific or standards consensus.
- The adoption profile, consequence class, jurisdiction, or data sensitivity changes.

Research conclusion

Research reports inform candidate architectures and controlled experiments. They do not create implementation assurance until their high-weight claims are reproduced and their limitations remain acceptable for the intended deployment profile.