M MYTHOS SYSTEMS

ENGINEERING STANDARDS LIBRARY / PORTFOLIO EDITION

Identity and Learning Standards Portfolio

Human, workload, agent, and service identity together with authorization, credentials, adaptive learning, and workforce assurance.

PERMANENT DOCUMENT ID
MYTHOS-IDEDU-PORT-0001

PUBLICATION
Portfolio Edition

STATUS
Candidate Portfolio

M MYTHOS SYSTEMS

ENGINEERING STANDARDS LIBRARY / ENGINEERING GUIDE

Knowledge, Data, Identity, and Learning Governance Guide

A permanent governance guide for source authority, state, evidence, identity, authorization, persistent memory, adaptive learning, and workforce assurance.

PERMANENT DOCUMENT ID
MYTHOS-GUIDE-KDIL-0001

PUBLICATION
Engineering Guide

STATUS
Candidate Guide

Contents

Executive operating doctrine

Knowledge, data, identity, and learning form the durable substrate around AI. Grounded decisions require authoritative sources, controlled state, explicit identity, verifiable evidence, user-governed memory, and measurable human capability.

Standards map

ID	PUBLICATION	CRITERIA
MYTHOS-KNW-0001	RAG, Grounding & Source Authority Standard Defines retrieval-augmented generation, grounding, citation, source authority, freshness, conflict handling, provenance, and evidence requirements for knowledge-backed AI systems.	12
MYTHOS-KNW-0002	Persona, Avatar & Persistent Memory Architecture Standard Establishes architecture and governance for persistent persona, embodied presence, user-controlled memory, continuity, provenance, privacy, and lifecycle management.	12
MYTHOS-DAT-0001	Vector Database & Local-First Sync Architecture Standard Defines vector retrieval and local-first synchronization architecture, including indexing, embeddings, consistency, conflict resolution, offline operation, provenance, and lifecycle control.	12
MYTHOS-DAT-0002	AI & Agent Observability Semantic Conventions (OpenTelemetry) Standard Establishes semantic conventions and operational requirements for observing models, agents, tools, retrieval, memory, costs, safety events, and end-to-end AI execution traces.	11
MYTHOS-DAT-0003	Data Sovereignty, Privacy Preservation, & Egress Control Standard Defines data sovereignty, privacy preservation, residency, classification, egress control, minimization, policy enforcement, and accountable cross-boundary processing.	14
MYTHOS-DAT-0004	Event Sourcing, CQRS, & Durable State Machine Standard Establishes event-sourcing, CQRS, durable state-machine, replay, idempotency, consistency, migration, and evidence requirements for authoritative state systems.	12
MYTHOS-DAT-0005	Tamper-Evident Hash-Chained Ledgers & Evidence Bundle Standard Defines tamper-evident ledgers and evidence bundles for consequential system actions, including hash chaining, signatures, provenance, retention, verification, and disclosure boundaries.	12
MYTHOS-ID-0001	Workload, Agent & Human Identity Standard Defines identity architecture for humans, workloads, services, agents, tools, and devices with attestation, federation, lifecycle, delegation, authentication, and audit requirements.	12
MYTHOS-ID-0002	Public Key Infrastructure (PKI) & Attribute-Based Access Control (ABAC) Standard Establishes public-key infrastructure and attribute-based access-control requirements for trust anchors, certificate lifecycle, policy decision, authorization context, revocation, and evidence.	12
MYTHOS-EDU-0001	Adaptive Learning, Skill Graphs & Cyber Lab Standard Defines adaptive learning, skill graphs, competency evidence, cyber labs, assessment, accessibility, credential portability, safety, and workforce-assurance requirements.	12

Connected governance chain

SYSTEM	AUTHORITATIVE QUESTION	CONTROL REQUIREMENT
Knowledge	What sources support the decision, how authoritative are they, and how fresh are they?	Preserve citations, source snapshots, conflicts, access decisions, and uncertainty.
Data and state	What is the authoritative state, who may change it, and how is it recovered?	Use typed state models, durable events, integrity, replay, retention, and reconciliation.
Identity	Which human, workload, service, agent, device, or tool is acting?	Authenticate and attest the principal; bind lifecycle, delegation, and revocation.
Authorization	What may this principal do to this resource in this context?	Evaluate policy externally and issue narrow, expiring, observable capability.
Evidence	How can the decision and resulting state be independently reconstructed?	Retain tamper-evident provenance, approvals, verification, and final disposition.
Learning	What capability changed, how was it measured, and what evidence supports the credential?	Use transparent skill models, authentic assessment, accessibility, and revocable credentials.

Knowledge authority and grounding

SOURCE CLASS	TYPICAL AUTHORITY	REQUIRED TREATMENT
Normative authority	Law, regulation, contract, approved policy, or controlled standard.	Pin exact version, jurisdiction, applicability, owner, and expiration.
Primary technical source	Official specification, standard, protocol registry, or product documentation.	Record version and retrieval date; distinguish specification from observed behavior.
Operational evidence	Logs, tests, measurements, tickets, state snapshots, and signed records.	Protect integrity, time, identity, collection method, retention, and access.
Qualified analysis	Reviewed research, expert assessment, or approved internal analysis.	Record method, assumptions, conflicts, confidence, and review date.
Secondary synthesis	Books, articles, summaries, and explanatory material.	Use for context, not as sole support for consequential claims.
Untrusted or user-generated	Forums, model output, anonymous material, and uncontrolled uploads.	Isolate, label, validate, and prohibit silent promotion into authoritative memory.

Data, state, and evidence architecture

STATE CONCERN	DESIGN QUESTION	ASSURANCE MECHANISM
Authority	Which store or event stream is authoritative for each entity and decision?	Explicit source-of-truth ownership and conflict policy.
Consistency	What stale, divergent, or partial states are acceptable?	Declared consistency model, version vectors, idempotency, and reconciliation.
Durability	What must survive process, host, region, and dependency failure?	Write-ahead or event records, checkpoints, backup, restore, and recovery testing.
Privacy and sovereignty	Where may data be stored, processed, embedded, retrieved, or disclosed?	Classification, residency, minimization, egress control, deletion, and legal basis.
Integrity	How are unauthorized change and evidence tampering detected?	Signatures, hashes, append-only records, access control, and verification.
Lifecycle	When is state corrected, migrated, archived, expired, or destroyed?	Versioned schemas, retention schedules, tombstones, and auditable deletion.
Observability	Can operators explain state transitions and diagnose divergence?	Correlated traces, events, metrics, logs, and state snapshots.

Identity, delegation, and authorization

PRINCIPAL	IDENTITY PROOF	AUTHORIZATION POSTURE
Human	Phishing-resistant authentication, account lifecycle, device and session context.	Role, attribute, risk, purpose, and step-up controls with accountable approval.
Workload or service	Workload identity, platform attestation, signed deployment, and runtime context.	Short-lived service capability bound to environment, audience, and resource.
Agent	Authenticated runtime instance, mission identity, model and configuration lineage.	Delegated capability restricted by mission, tool, target, budget, and expiration.
Tool or extension	Signed publisher identity, package provenance, version, permissions, and trust tier.	Explicit grants, sandbox tier, network and filesystem restrictions, and revocation.
Device	Hardware or software identity, posture, ownership, and attestation.	Conditional access based on compliance, location, risk, and intended function.
Data subject or persona	Verified account relationship, consent, preferences, and representation constraints.	Purpose limitation, user control, privacy, correction, deletion, and impersonation resistance.

Adaptive learning and workforce assurance

LEARNING ELEMENT	CONTROLLED IMPLEMENTATION	EVIDENCE
Skill graph	Versioned competencies, prerequisites, proficiency levels, and role mappings.	Approved ontology and change history.
Adaptive path	Recommendations based on demonstrated gaps, goals, accessibility, and pace.	Explainable recommendation and learner control.
Cyber lab	Isolated, resettable, observable environment with scenario and safety boundaries.	Lab configuration, activity telemetry, and cleanup evidence.
Assessment	Authentic task, rubric, anti-cheating controls, and human escalation.	Scored artifact, evaluator identity, method, and confidence.
Credential	Portable, verifiable achievement tied to evidence and issuer.	Signed credential, criteria, evidence reference, issue and expiration.
Continuous assurance	Revalidation after time, role, technology, or risk change.	Review schedule, reassessment result, remediation, and revocation state.

Cross-domain failure and recovery

FAILURE	IMPACT	CONTROL RESPONSE
Source-authority inversion	Low-quality or adversarial content becomes trusted knowledge.	Quarantine, compare authority, restore lineage, and invalidate derived memory.
Stale grounding	Decision cites an obsolete specification, policy, identity, or state snapshot.	Enforce freshness windows, expiration, revalidation, and conflict display.
Identity confusion	Agent, service, person, or device is misattributed.	Deny action, reauthenticate or attest, repair correlation, and review delegated capabilities.
Excessive delegation	A principal transfers broader or longer authority than intended.	Revoke grants, narrow policy, require approval, and inspect prior activity.
State divergence	Offline, distributed, or cached copies disagree with authority.	Enter reconciliation, preserve conflicting versions, and prevent unsafe promotion.
Evidence break	Records cannot prove sequence, identity, integrity, or final state.	Classify result as unassured and restore instrumentation before resuming.
Learning manipulation	Assessment, skill graph, or adaptive path is gamed or biased.	Review evidence, recalibrate, require alternate assessment, and provide appeal.
Privacy over-retention	Memory, embeddings, traces, or credentials persist beyond purpose.	Delete, revoke, rebuild derived indexes, and document residual copies.

Minimum evidence by decision domain

DECISION DOMAIN	MINIMUM EVIDENCE
Grounded answer	Claim-to-source mapping, source authority, freshness, retrieval trace, conflicts, and confidence.
State transition	Prior state, command or event, authenticated principal, policy decision, resulting state, and verification.
Identity issuance	Proofing method, authenticator or workload attestation, issuer, policy, validity, and revocation path.
Authorization	Principal, attributes, resource, action, context, policy version, decision, obligations, and expiration.
Memory write	Source, purpose, scope, confidence, sensitivity, owner, expiration, correction, and deletion semantics.
Credential award	Competency, rubric, assessment evidence, evaluator, issuer, issue date, validity, and revocation.
Data movement	Classification, source, destination, purpose, legal or policy basis, transformation, egress decision, and retention.
Exception	Control, rationale, risk, compensating controls, owner, approval, expiration, and review result.

Implementation roadmap

PHASE	FOCUS	EXIT CONDITION
1. Authority model	Source classes, ownership, exact versions, freshness, and conflict policy.	Approved authority registry and review cadence.
2. Identity substrate	Human, workload, agent, device, and extension identity with lifecycle and attestation.	End-to-end principal correlation and revocation demonstration.
3. State and evidence	Authoritative models, durable events, integrity, replay, and evidence bundles.	Recovery and reconstruction tests pass.
4. Knowledge and memory	Grounded retrieval, provenance, memory scopes, correction, deletion, and contradiction handling.	Poisoning, stale-source, and privacy tests pass.
5. Learning assurance	Skill graphs, accessible labs, authentic assessments, and verifiable credentials.	Evidence-backed credential issuance and appeal process.
6. Continuous governance	Metrics, incidents, exceptions, audits, source expiration, and revalidation.	Operating owners accept SLOs and review obligations.

Assurance conclusion

Knowledge informs decisions; identity establishes the actor; policy constrains authority; data records state; evidence makes outcomes reviewable; learning changes capability under transparent and measurable governance. Weakness in any one layer limits the assurance of the entire system.

ENGINEERING STANDARDS LIBRARY / IDENTITY AND ACCESS

Workload, Agent & Human Identity Standard

The architecture of identity has failed.

PERMANENT DOCUMENT ID

MYTHOS-ID-0001

PUBLICATION

Candidate Standard
Version 1.0.0-candidate

ASSURANCE PROFILE

Standard / Advanced
High Assurance

PERMANENT PUBLICATION IDENTITY

Workload, Agent & Human Identity Standard

DOCUMENT ID
MYTHOS-ID-0001

VERSION
1.0.0-candidate

STATUS
Candidate Standard

PUBLISHER
Mythos Systems

PUBLICATION DATE
2026-07-24

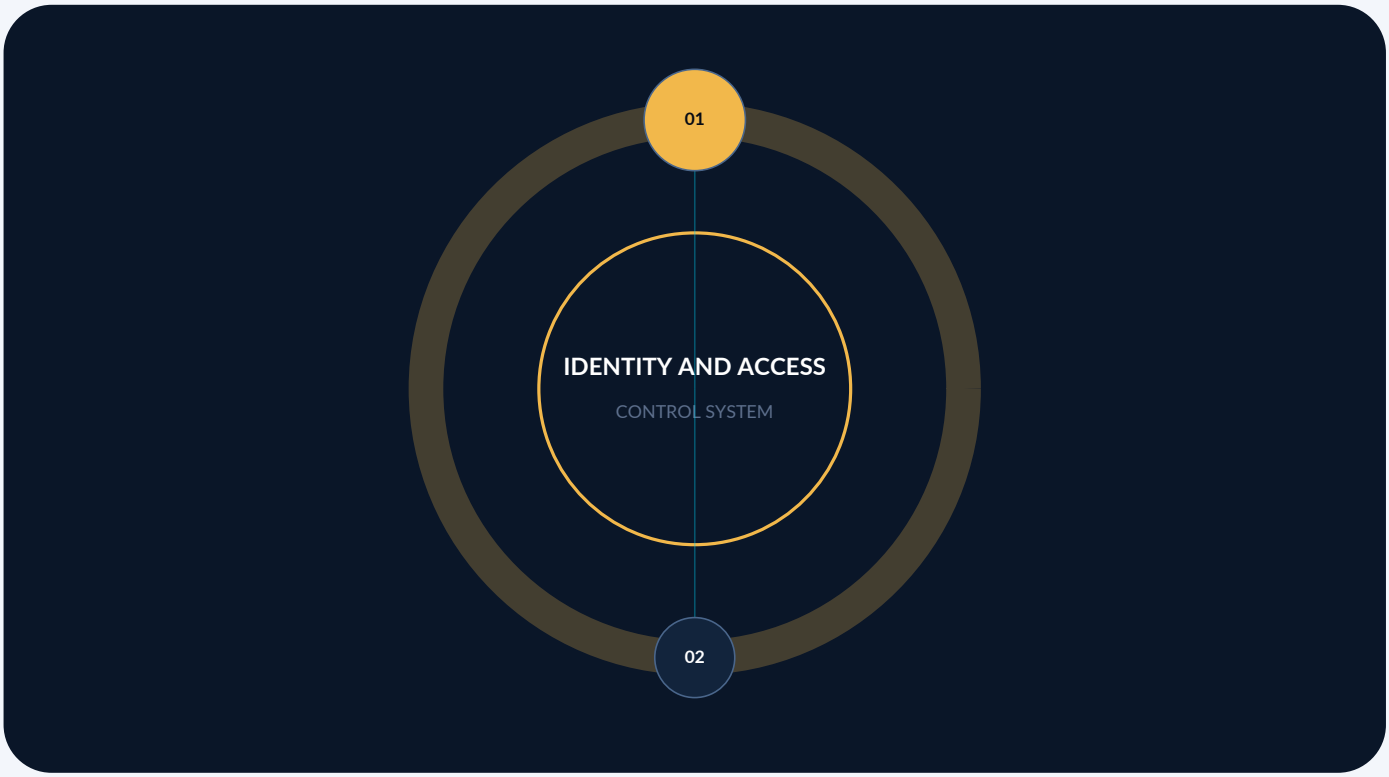
SCHEDULED REVIEW
2027-07-24

12	6	4	1
ATOMIC CRITERIA	ASSURANCE VIEWS	IMPLEMENTATION PROFILES	PERMANENT IDENTITY

Publication doctrine. This standard is a permanent library identity. Interpretation must preserve its recorded evidence, controlled exceptions, formal revision history, and explicit relationship to companion standards.

DOMAIN CONTROL SYSTEM

MYTHOS-ID-0001 inside the identity and access constellation

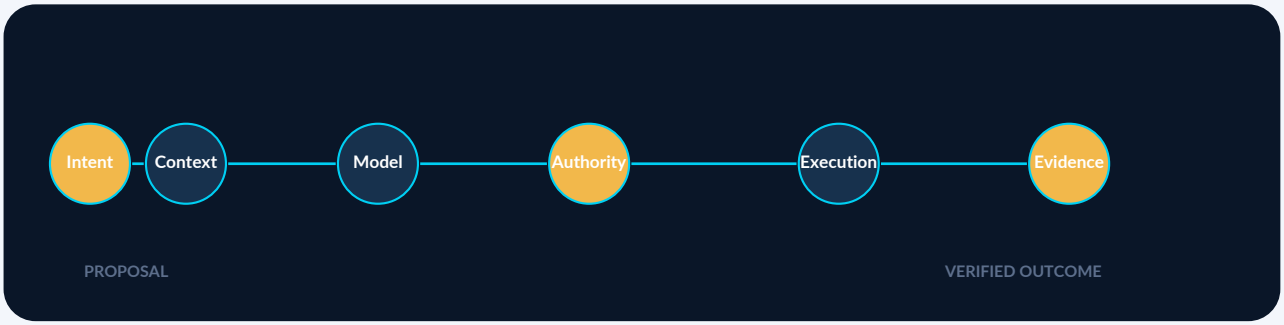


Connected assurance	Architecture, implementation, operations, evidence, and lifecycle are evaluated as one controlled system.
Specialization rule	Companion standards may add precision, but cannot silently weaken this publication.

Executive orientation

Defines identity architecture for humans, workloads, services, agents, tools, and devices with attestation, federation, lifecycle, delegation, authentication, and audit requirements.

WHY THIS STANDARD EXISTS	The architecture of identity has failed.
OPERATIONAL SCOPE	This standard covers the evaluation of: - Human identity: Passkeys (FIDO2, WebAuthn, Apple/Google/Microsoft sync) - Workload identity: SPIFFE/SPIRE, mTLS, cloud IAM instance profiles - Agent identity: Ephemeral tokens, Macaroons, capability-based delegation - Authorization protocols: OAuth 2.1, OIDC, token exchange - Privileged Access Management (PAM): Just-in-Time (JIT) access, step-up auth - Cryptographic token lifecycle: Issuance, validation, rotation, revocation
PRIMARY AUDIENCE	- Security architects and IAM engineers - Platform engineers building zero-trust infrastructure - AI engineers building agent authentication and delegation - Compliance and audit teams managing access governance - Standards bodies seeking a reference identity methodology



Assurance principle. Model capability, data quality, identity, authority, operational evidence, and human accountability are treated as a connected system. A high aggregate score cannot compensate for a failed mandatory safety or authority boundary.

Contents

Evaluation criterion index

This standard contains 12 atomic criteria. Each criterion is independently testable and requires retained evidence.

ID	EVALUATION CRITERION
4.1.1	Phishing Resistance
4.2.1	Machine Secret Elimination
4.3.1	Authority Scoping
4.4.1	Standing Privilege Elimination
4.5.1	Token Validity
4.6.1	Identity Observability
5.1.1	M-ID-S-Human
5.1.2	M-ID-S-Agent
6.1.1	Adversary-in-the-Middle (AiTM) Phishing
6.1.2	Agent Token Hijacking
6.1.3	Standing Privilege Abuse
6.1.4	Confused Deputy (Scope Creep)

MYTHOS-ID-0001-4.1.1

4.1.1 Phishing Resistance

Normative source requirement

Are human users authenticated via cryptographic possession rather than shared secrets?

Score	Criteria
0	Passwords only
1	Passwords + SMS/TOTP MFA (phishable)
2	Push-based MFA (phishable via MFA fatigue)
3	FIDO2/WebAuthn passkeys (phishing-resistant)
4	Device-bound passkeys requiring local biometric user verification for every auth
5	Perfect resistance: formally verified key bounds, hardware-bound nonces, zero shared secrets

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-4.2.1

4.2.1 Machine Secret Elimination

Normative source requirement

How do microservices and workloads authenticate to each other?

Score	Criteria
0	Static API keys, passwords, or cloud IAM keys in environment variables
1	Cloud-managed instance profiles (AWS IAM Roles for Service Accounts)
2	Internal PKI with long-lived certificates
3	SPIFFE/SPIRE issuing short-lived SVIDs (SPIFFE Verifiable Identity Documents)
4	Automated SVID rotation (< 5-minute TTLs); zero static secrets in the environment
5	Perfect identity: formally verified workload attestation, zero-trust mTLS mesh, cryptographic proof of workload binary signature before SVID issuance

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-4.3.1

4.3.1 Authority Scoping

Normative source requirement

How is an autonomous agent's authority constrained when acting on behalf of a user?

Score	Criteria
0	Agent is passed the user's full OAuth token or password
1	Agent uses a generic service account with broad API access
2	Agent uses OAuth token exchange (RFC 8693) to reduce scopes, but token is long-lived
3	Ephemeral agent identity: short-lived token bound to a specific task and user
4	Cryptographic attenuation: agent cannot delegate more authority than it possesses (Macaroons/Biscuits)
5	Perfect delegation: formally verified capability chains, zero scope expansion possible, cryptographic proof of user intent bound to agent action

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-4.4.1

4.4.1 Standing Privilege Elimination

Normative source requirement

Are administrative rights granted continuously or only when needed?

Score	Criteria
0	Users have persistent "admin" or "root" roles
1	Break-glass accounts exist, but manual checkout process
2	PAM solution used to rotate passwords for standing admin accounts
3	Just-in-Time (JIT) access: zero standing privilege; rights granted dynamically via approval
4	JIT access with automatic expiration and real-time anomaly detection on session behavior
5	Perfect privilege: formally verified zero-standing access, sub-minute provisioning, cryptographic attestation of PAM session boundaries

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-4.5.1

4.5.1 Token Validity

Normative source requirement

How are authorization tokens issued, validated, and revoked?

Score	Criteria
0	Long-lived JWTs with no expiration or revocation mechanism
1	Short-lived JWTs, but no revocation (must wait for expiry)
2	Refresh tokens with immediate revocation capability
3	OAuth 2.1 DPoP (Demonstrating Proof-of-Possession) binding tokens to client keys
4	Transaction-bound tokens: token contains a hash of the specific action it authorizes
5	Perfect lifecycle: formally verified token bounds, zero replay capability, cryptographic non-repudiation of token use

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-4.6.1

4.6.1 Identity Observability

Normative source requirement

Can security teams reconstruct the complete chain of identity and delegation?

Score	Criteria
0	Logs only record "user X logged in"
1	Logs record API calls, but no token context
2	Logs link actions to specific OAuth tokens
3	Full audit trail: user -> OIDC -> token -> delegation -> agent -> workload
4	Real-time anomaly detection on identity events (impossible travel, token reuse)
5	Perfect observability: formally verified identity provenance graphs, cryptographic signing of all identity events, zero blind spots in delegation chains

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-5.1.1

5.1.1 M-ID-S-Human

Normative source requirement

- Phishing Simulation: 100+ tests attempting to intercept credentials via reverse proxies, verifying Passkeys/FIDO2 resist the relay.
- MFA Fatigue: 50+ tests spamming push notifications, verifying the system requires transaction-bound user presence.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Architecture records, implemented configuration, test evidence, operational telemetry, exception decisions, and accountable approval.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-5.1.2

5.1.2 M-ID-S-Agent

Normative source requirement

• Scope Escalation: 100+ tests attempting to use an agent's delegated token to access unauthorized APIs, verifying the capability scoping blocks the action.

• Token Replay: 50+ tests attempting to reuse an expired or already-consumed agent token.

• Attenuation Break: 50+ tests attempting to forge a Macaroon/capability token to gain higher privileges, verifying the cryptographic signature fails.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Trace the decision path from identity and policy through execution, attempt bypass and privilege-escalation scenarios, verify revocation behavior, and inspect the resulting evidence record.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-6.1.1

6.1.1 Adversary-in-the-Middle (AiTM) Phishing

Normative source requirement

Severity: Critical Description: An attacker uses a reverse proxy (e.g., Evilginx) to steal session cookies and bypass traditional MFA.

Required Controls: 1. FIDO2/WebAuthn Passkeys. The cryptographic challenge is origin-bound, meaning a proxy cannot relay the response to the legitimate backend.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Architecture records, implemented configuration, test evidence, operational telemetry, exception decisions, and accountable approval.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-6.1.2

6.1.2 Agent Token Hijacking

Normative source requirement

Severity: Critical Description: A prompt injection or memory poisoning attack causes an agent to exfiltrate its delegated OAuth token to an attacker-controlled server.

Required Controls: 1. Tokens MUST be short-lived (sub-15 minutes). 2. Tokens MUST be bound to the specific task (transaction-bound), rendering them useless to the attacker. 3. Where possible, use DPoP so the token is bound to the agent's private key, which the attacker does not possess.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Context manifests, retrieval traces, source citations, freshness records, conflict decisions, and memory provenance.
ASSESSMENT METHOD	MATURITY SIGNAL
Trace the decision path from identity and policy through execution, attempt bypass and privilege-escalation scenarios, verify revocation behavior, and inspect the resulting evidence record.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-6.1.3

6.1.3 Standing Privilege Abuse

Normative source requirement

Severity: High Description: A compromised admin account is used to create backdoors or exfiltrate data because the account had continuous root access.

Required Controls: 1. Just-in-Time (JIT) PAM. The account has zero standing privilege and must request elevation, triggering an approval workflow and audit trail.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Trace the decision path from identity and policy through execution, attempt bypass and privilege-escalation scenarios, verify revocation behavior, and inspect the resulting evidence record.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0001-6.1.4

6.1.4 Confused Deputy (Scope Creep)

Normative source requirement

Severity: High Description: An agent is granted an OAuth scope to "read files," but the API implicitly allows it to "delete files" because the scope was too broad.

Required Controls: 1. Granular capability-based scoping (Macaroons). 2. The token must explicitly state the exact action (`action:read`) and target (`file:123`).

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Architecture records, implemented configuration, test evidence, operational telemetry, exception decisions, and accountable approval.
ASSESSMENT METHOD	MATURITY SIGNAL
Exercise crash, retry, duplicate, reorder, partition, and restore scenarios; compare authoritative state before and after replay; confirm idempotency and recovery evidence.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

Current authority and freshness register

These sources inform interpretation but do not convert this candidate publication into certification, legal compliance, or implementation evidence. Rolling sources must be version-pinned at adoption time.

AUTHORITY	VERSION / FRESHNESS	APPLICABILITY LIMIT
NIST - Digital Identity Guidelines	SP 800-63-4 final, July 2025 Expires 2027-01-24	Federal guidance must be tailored for non-federal risk, jurisdiction, usability, privacy, and operational context.
W3C - Verifiable Credentials Data Model v2.0	W3C Recommendation, May 15 2025 Expires 2027-01-24	Data model conformance does not establish issuer trust, credential truth, or ecosystem governance.
W3C - Decentralized Identifiers (DIDs) v1.0	W3C Recommendation; DID 1.1 remains a Candidate Recommendation Expires 2027-01-24	DID method security, governance, resolution, and privacy properties vary materially.
SPIFFE - SPIFFE Specifications and Workload API	Rolling specification snapshot; SPIRE 1.15.2 documentation Expires 2026-10-24	Implementation and deployment assurance depend on attestation plugins, trust-domain design, and operational controls.

Source lineage appendix

The following text preserves the substantive source draft in a normalized public form. Where it conflicts with the permanent publication identity, candidate status, current authority register, or evidence requirements above, the controlled publication layer governs.

0. Executive Mandate

The architecture of identity has failed.

Organizations secure multi-cloud architectures and autonomous agent fleets using static API keys, shared passwords, and long-lived service accounts. When a secret is compromised, it is manually rotated. When an agent acts on behalf of a user, it inherits the user's entire scope of authority with no mechanism for attenuation. This is not security; it is accountability theater.

This standard exists because:

1. Passwords and Static Keys are Architectural Liabilities: A shared secret is a single point of failure. If a secret is stolen, the attacker becomes indistinguishable from the legitimate user. Human identity MUST be based on phishing-resistant, asymmetric cryptography (FIDO2/WebAuthn Passkeys). Machine identity MUST be based on ambient, short-lived cryptographic identity (SPIFFE).

2. Agents are Not Humans, Nor are They Service Accounts: An autonomous agent is a new class of entity. It requires an ephemeral identity that is bound to a specific task, a specific user delegation, and a strict time-to-live (TTL). Agents MUST NOT inherit blanket user credentials; they MUST operate via cryptographically attenuated delegation chains.

3. Standing Privilege is an Attack Vector: A user or agent with continuous "admin" access is a compromised asset waiting to happen. Privileged Access Management (PAM) MUST enforce Just-in-Time (JIT) access, elevating privileges only for the duration of a specific, approved task.

4. OAuth 2.0 is Insufficient Without Strict Scoping: Traditional OAuth often results in "scope creep," where a token grants access to an entire API rather than a specific action. Token issuance MUST be bound to granular capabilities, and high-risk actions MUST require transaction-bound step-up authentication.

This document establishes the Mythos Identity Standard (M-ID-S), a comprehensive, evidence-based methodology for evaluating identity architectures against zero-trust principles, cryptographic non-repudiation, and ephemeral authority.

Part I. Scope and Applicability

1.1 In Scope

This standard covers the evaluation of:

- Human identity: Passkeys (FIDO2, WebAuthn, Apple/Google/Microsoft sync)
- Workload identity: SPIFFE/SPIRE, mTLS, cloud IAM instance profiles
- Agent identity: Ephemeral tokens, Macaroons, capability-based delegation
- Authorization protocols: OAuth 2.1, OIDC, token exchange
- Privileged Access Management (PAM): Just-in-Time (JIT) access, step-up auth
- Cryptographic token lifecycle: Issuance, validation, rotation, revocation

1.2 Out of Scope

- General zero-trust network architecture (covered in MYTHOS-NET-0004)
- General HITL workflow design (covered in MYTHOS-UX-0002)
- Post-Quantum Cryptography algorithms (covered in MYTHOS-SEC-0001)

1.3 Audience

- Security architects and IAM engineers
- Platform engineers building zero-trust infrastructure
- AI engineers building agent authentication and delegation
- Compliance and audit teams managing access governance
- Standards bodies seeking a reference identity methodology

Part II. Definitions and Taxonomy

2.1 Identity Dimensions

Dimension	Definition
Passkey	A phishing-resistant credential based on FIDO2/WebAuthn, utilizing public-key cryptography where the private key is bound to a hardware secure enclave or synced OS keychain.
SPIFFE	Secure Production Identity Framework for Everyone. A framework for issuing cryptographically verifiable identity documents (SVIDs) to workloads, eliminating the need for static secrets.
Ephemeral Agent Identity	A short-lived, cryptographically signed identity token issued to an autonomous agent, bound to a specific user delegation and task scope, expiring automatically upon completion.
Attenuation	The process of reducing the scope or authority of a token when it is delegated. A child agent cannot have more authority than the parent agent or delegating user.
Just-in-Time (JIT) PAM	An access model where users/agents have zero standing privileges and must request and be granted elevated access for a specific, time-boxed window.

2.2 The Identity Doctrine

> A static secret is a liability. A password is a shared key waiting to be leaked. An agent that holds a user's blanket token is an unconstrained delegation. If the privilege is standing, the breach is already underway.

Part III. Evaluation Methodology

3.1 Scoring Model

Each capability is scored from 0 to 5.

Score	Meaning
0	Fails basic task; passwords, static API keys, standing privileges
1	Basic MFA and service accounts, but no workload identity or agent scoping
2	Functional OAuth/OIDC, but relies on long-lived tokens or manual PAM
3	Solid production performance; Passkeys, SPIFFE, JIT PAM, token attenuation
4	Strong performance; transaction-bound step-up auth, zero-standing privilege
5	Best-in-class; sets the standard for mathematically verified, zero-trust identity

Weighting

Category	Weight	Rationale
Human Authentication (Passkeys)	15%	Passwords are the #1 initial access vector
Workload Identity (SPIFFE/mTLS)	20%	Static secrets in CI/CD and microservices cause cascading breaches
Agent Delegation & Attenuation	25%	Agents are the new perimeter; unscoped agents are catastrophic
Privileged Access (PAM/JIT)	15%	Standing admin privileges guarantee total system compromise
Token Lifecycle & Cryptography	15%	Long-lived tokens are stolen tokens
Operational Lifecycle & Auditing	10%	Identity must be observable and verifiable

Total: 100%

A system MUST score at least 3.0 in Workload Identity and Agent Delegation to be considered for production use.

Part IV. Evaluation Categories

4.1 Human Authentication (Passkeys) (Weight: 15%)

4.1.1 Phishing Resistance

Are human users authenticated via cryptographic possession rather than shared secrets?

Score	Criteria
0	Passwords only
1	Passwords + SMS/TOTP MFA (phishable)
2	Push-based MFA (phishable via MFA fatigue)
3	FIDO2/WebAuthn passkeys (phishing-resistant)
4	Device-bound passkeys requiring local biometric user verification for every auth
5	Perfect resistance: formally verified key bounds, hardware-bound nonces, zero shared secrets

4.2 Workload Identity (SPIFFE/mTLS) (Weight: 20%)

4.2.1 Machine Secret Elimination

How do microservices and workloads authenticate to each other?

Score	Criteria
0	Static API keys, passwords, or cloud IAM keys in environment variables
1	Cloud-managed instance profiles (AWS IAM Roles for Service Accounts)
2	Internal PKI with long-lived certificates
3	SPIFFE/SPIRE issuing short-lived SVIDs (SPIFFE Verifiable Identity Documents)
4	Automated SVID rotation (< 5-minute TTLs); zero static secrets in the environment
5	Perfect identity: formally verified workload attestation, zero-trust mTLS mesh, cryptographic proof of workload binary signature before SVID issuance

4.3 Agent Delegation and Attenuation (Weight: 25%)

4.3.1 Authority Scoping

How is an autonomous agent's authority constrained when acting on behalf of a user?

Score	Criteria
0	Agent is passed the user's full OAuth token or password
1	Agent uses a generic service account with broad API access
2	Agent uses OAuth token exchange (RFC 8693) to reduce scopes, but token is long-lived
3	Ephemeral agent identity: short-lived token bound to a specific task and user
4	Cryptographic attenuation: agent cannot delegate more authority than it possesses (Macaroons/Biscuits)

Score	Criteria
5	Perfect delegation: formally verified capability chains, zero scope expansion possible, cryptographic proof of user intent bound to agent action

4.4 Privileged Access (PAM/JIT) (Weight: 15%)

4.4.1 Standing Privilege Elimination

Are administrative rights granted continuously or only when needed?

Score	Criteria
0	Users have persistent "admin" or "root" roles
1	Break-glass accounts exist, but manual checkout process
2	PAM solution used to rotate passwords for standing admin accounts
3	Just-in-Time (JIT) access: zero standing privilege; rights granted dynamically via approval
4	JIT access with automatic expiration and real-time anomaly detection on session behavior
5	Perfect privilege: formally verified zero-standing access, sub-minute provisioning, cryptographic attestation of PAM session boundaries

4.5 Token Lifecycle and Cryptography (Weight: 15%)

4.5.1 Token Validity

How are authorization tokens issued, validated, and revoked?

Score	Criteria
0	Long-lived JWTs with no expiration or revocation mechanism
1	Short-lived JWTs, but no revocation (must wait for expiry)
2	Refresh tokens with immediate revocation capability
3	OAuth 2.1 DPoP (Demonstrating Proof-of-Possession) binding tokens to client keys
4	Transaction-bound tokens: token contains a hash of the specific action it authorizes
5	Perfect lifecycle: formally verified token bounds, zero replay capability, cryptographic non-repudiation of token use

4.6 Operational Lifecycle and Auditing (Weight: 10%)

4.6.1 Identity Observability

Can security teams reconstruct the complete chain of identity and delegation?

Score	Criteria
0	Logs only record "user X logged in"
1	Logs record API calls, but no token context
2	Logs link actions to specific OAuth tokens
3	Full audit trail: user -> OIDC -> token -> delegation -> agent -> workload
4	Real-time anomaly detection on identity events (impossible travel, token reuse)
5	Perfect observability: formally verified identity provenance graphs, cryptographic signing of all identity events, zero blind spots in delegation chains

Part V. The Mythos Identity Suite (M-ID-S)

Traditional IAM benchmarks measure authentication latency. The M-ID-S evaluates phishing resistance, workload secret elimination, and agent delegation bounds.

5.1 Suite Composition

5.1.1 M-ID-S-Human

- Phishing Simulation: 100+ tests attempting to intercept credentials via reverse proxies, verifying Passkeys/FIDO2 resist the relay.
- MFA Fatigue: 50+ tests spamming push notifications, verifying the system requires transaction-bound user presence.

5.1.2 M-ID-S-Agent

- Scope Escalation: 100+ tests attempting to use an agent's delegated token to access unauthorized APIs, verifying the capability scoping blocks the action.
- Token Replay: 50+ tests attempting to reuse an expired or already-consumed agent token.
- Attenuation Break: 50+ tests attempting to forge a Macaroon/capability token to gain higher privileges, verifying the cryptographic signature fails.

5.2 Execution Protocol

```
1. Deploy IAM architecture with Passkeys, SPIFFE, and Agent delegation
2. Execute complex, multi-step agentic workflows requiring JIT access
3. Run M-ID-S suite (automated phishing + token replay + escalation)
4. Score each category 0-5
5. Check for static API keys in code or long-lived admin roles (instant disqualification)
6. If all thresholds met: approve for production
```

Part VI. Security Threat Model for Identity

6.1 Threat Catalog

6.1.1 Adversary-in-the-Middle (AiTM) Phishing

Severity: Critical Description: An attacker uses a reverse proxy (e.g., Evilginx) to steal session cookies and bypass traditional MFA.

Required Controls:

1. FIDO2/WebAuthn Passkeys. The cryptographic challenge is origin-bound, meaning a proxy cannot relay the response to the legitimate backend.

6.1.2 Agent Token Hijacking

Severity: Critical Description: A prompt injection or memory poisoning attack causes an agent to exfiltrate its delegated OAuth token to an attacker-controlled server.

Required Controls:

1. Tokens MUST be short-lived (sub-15 minutes).

2. Tokens MUST be bound to the specific task (transaction-bound), rendering them useless to the attacker.

3. Where possible, use DPoP so the token is bound to the agent's private key, which the attacker does not possess.

6.1.3 Standing Privilege Abuse

Severity: High Description: A compromised admin account is used to create backdoors or exfiltrate data because the account had continuous root access.

Required Controls:

1. Just-in-Time (JIT) PAM. The account has zero standing privilege and must request elevation, triggering an approval workflow and audit trail.

6.1.4 Confused Deputy (Scope Creep)

Severity: High Description: An agent is granted an OAuth scope to "read files," but the API implicitly allows it to "delete files" because the scope was too broad.

Required Controls:

1. Granular capability-based scoping (Macaroons).

2. The token must explicitly state the exact action (`action:read`) and target (`file:123`).

Part VII. The Mythos Identity Standard

7.1 The Mythos Identity Doctrine

```
A static secret is a liability.
A password is a shared key waiting to be leaked.

An agent that holds a user's blanket token is an unconstrained delegation.
If the privilege is standing, the breach is already underway.

Humans may be unpredictable.
Identity must be absolute.
```

7.2 Selection Rules

1. No identity architecture enters production without passing M-ID-S.
2. No architecture is approved if it relies on passwords for human authentication.
3. No architecture is approved if workloads use static API keys.
4. No agent is approved if it holds a user's persistent OAuth token.
5. No architecture is approved with standing administrative privileges.

7.3 The Prime Directive for Identity Architecture

> Bind the key to the hardware, do not share the secret. > Issue the SVID to the workload, do not embed the credential. > Attenuate the agent token, do not delegate the blanket scope. > Expire the privilege, do not grant the standing role.

```
Academic benchmarks measure authentication latency.
Applied benchmarks measure workload secret elimination.
Security benchmarks measure phishing resistance.
Operational benchmarks measure delegation attenuation.

> Identity may be federated.
> Authority must be absolute.
```

End of Standard MYTHOS-ID-0001

End of controlled publication

MYTHOS-ID-0001 remains a Candidate Standard until technical review, security and privacy review, accessibility review, current-source validation, and formal approval are recorded.

ENGINEERING STANDARDS LIBRARY / IDENTITY AND ACCESS

Public Key Infrastructure (PKI) & Attribute-Based Access Control (ABAC) Standard

The architecture of authorization and key management has failed.

PERMANENT DOCUMENT ID

MYTHOS-ID-0002

PUBLICATION

Candidate Standard
Version 1.0.0-candidate

ASSURANCE PROFILE

Standard / Advanced
High Assurance

PERMANENT PUBLICATION IDENTITY

Public Key Infrastructure (PKI) & Attribute-Based Access Control (ABAC) Standard

DOCUMENT ID
MYTHOS-ID-0002

VERSION
1.0.0-candidate

STATUS
Candidate Standard

PUBLISHER
Mythos Systems

PUBLICATION DATE
2026-07-24

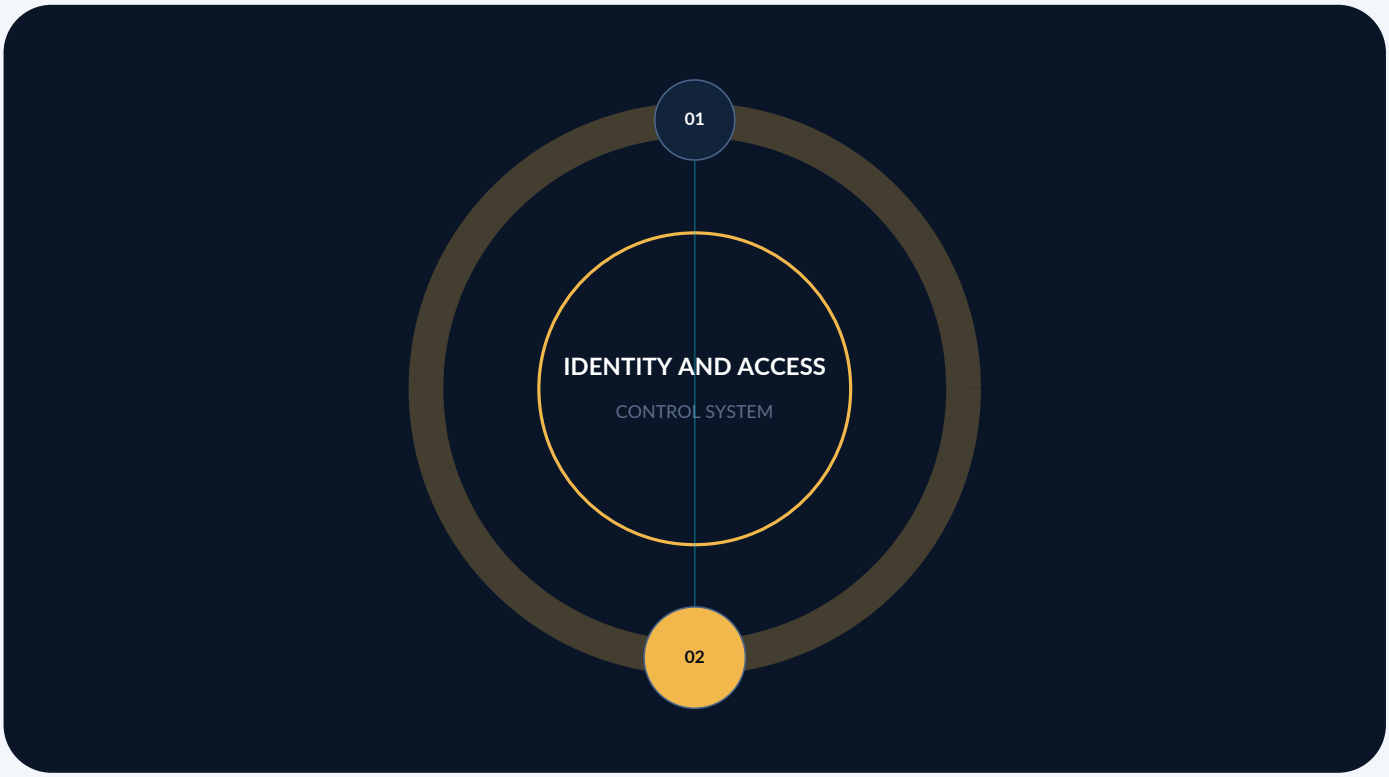
SCHEDULED REVIEW
2027-07-24

12	6	4	1
ATOMIC CRITERIA	ASSURANCE VIEWS	IMPLEMENTATION PROFILES	PERMANENT IDENTITY

Publication doctrine. This standard is a permanent library identity. Interpretation must preserve its recorded evidence, controlled exceptions, formal revision history, and explicit relationship to companion standards.

DOMAIN CONTROL SYSTEM

MYTHOS-ID-0002 inside the identity and access constellation

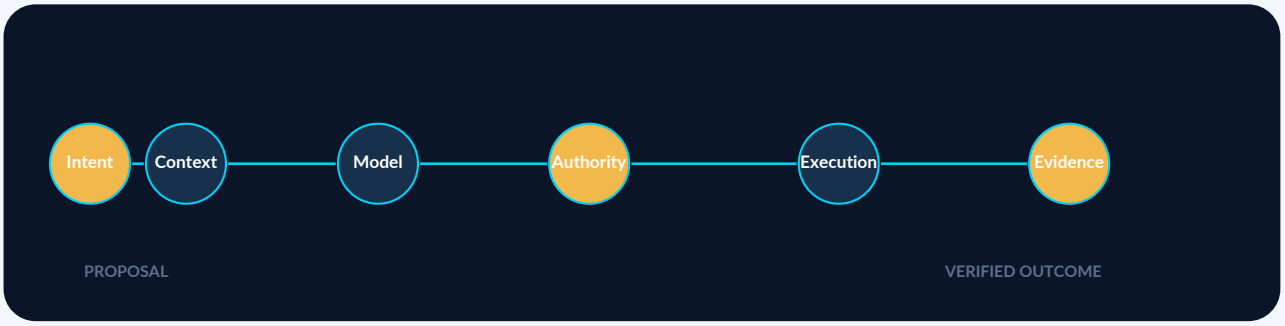


Connected assurance	Architecture, implementation, operations, evidence, and lifecycle are evaluated as one controlled system.
Specialization rule	Companion standards may add precision, but cannot silently weaken this publication.

Executive orientation

Establishes public-key infrastructure and attribute-based access-control requirements for trust anchors, certificate lifecycle, policy decision, authorization context, revocation, and evidence.

WHY THIS STANDARD EXISTS	The architecture of authorization and key management has failed.
OPERATIONAL SCOPE	This standard covers the evaluation of: - Public Key Infrastructure (PKI) and Certificate Authority (CA) hierarchies - Certificate lifecycle automation (ACME, cert-manager, SPIFFE SVIDs) - Hardware Security Modules (HSM) and TPM integration - Attribute-Based Access Control (ABAC) and Policy-as-Code (Open Policy Agent, Cedar) - The transition from Role-Based Access Control (RBAC) to ABAC - mTLS (Mutual TLS) enforcement and short-lived identity documents
PRIMARY AUDIENCE	- Security architects and PKI engineers - Platform engineers building zero-trust service meshes - Application architects decoupling authorization logic from microservices - Compliance teams managing access governance - Standards bodies seeking a reference PKI and ABAC methodology



Assurance principle. Model capability, data quality, identity, authority, operational evidence, and human accountability are treated as a connected system. A high aggregate score cannot compensate for a failed mandatory safety or authority boundary.

Contents

Evaluation criterion index

This standard contains 12 atomic criteria. Each criterion is independently testable and requires retained evidence.

ID	EVALUATION CRITERION
4.1.1	PKI Operations
4.2.1	Root of Trust
4.3.1	Authorization Model
4.4.1	Enforcement Architecture
4.5.1	Data Inputs
4.6.1	Compliance Evidence
5.1.1	MPAS-PKI
5.1.2	MPAS-ABAC
6.1.1	Certificate Authority Compromise
6.1.2	Role Explosion (RBAC Privilege Creep)
6.1.3	Stale Policy Enforcement
6.1.4	Certificate Expiry Outage

MYTHOS-ID-0002-4.1.1

4.1.1 PKI Operations

Normative source requirement

How are X.509 certificates issued, renewed, and revoked?

Score	Criteria
0	Manual CSRs; certificates last years; tracked in spreadsheets
1	Internal CA exists, but clients must manually request and install certs
2	cert-manager used in Kubernetes, but other infrastructure is manual
3	Full ACME integration across all workloads (on-prem, cloud, edge); certificates last < 90 days
4	Ephemeral identity (SPIFFE SVIDs) with cert TTLs < 1 hour; zero manual rotation
5	Perfect automation: formally verified lifecycle bounds, zero certificate-induced outages, mathematical proof of continuous validity

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-4.2.1

4.2.1 Root of Trust

Normative source requirement

Are Certificate Authority private keys protected against exfiltration?

Score	Criteria
0	Root CA private key stored on a standard VM or in plaintext
1	Intermediate CA keys password-protected on disk
2	Cloud-managed HSM (AWS CloudHSM, Azure Key Vault Premium) used for CAs
3	Dedicated, on-premises FIPS 140-2 Level 3+ HSMs; keys are non-exportable
4	TPM 2.0 attestation required for all leaf certificates; keys generated in silicon
5	Perfect security: formally verified hardware boundaries, zero ability to extract CA keys, cryptographic proof of key provenance

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-4.3.1

4.3.1 Authorization Model

Normative source requirement

Are access decisions static (based on groups) or dynamic (based on attributes)?

Score	Criteria
0	Pure RBAC; users assigned to groups ("Admins", "Users") with broad permissions
1	RBAC with some coarse context (e.g., IP-based restrictions)
2	Hybrid model: RBAC for broad access, ABAC for sensitive resources
3	Pure ABAC: roles abolished; access evaluated based on user attributes, resource tags, and environment
4	Real-time attribute evaluation: access changes dynamically if device posture degrades or geolocation shifts
5	Perfect model: formally verified policy bounds, zero standing privilege, mathematical proof of least-privilege enforcement

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-4.4.1

4.4.1 Enforcement Architecture

Normative source requirement

Is authorization logic decoupled from the application codebase?

Score	Criteria
0	Authorization logic hardcoded in application controllers (e.g., `if user.role == 'admin'`)
1	Centralized IAM, but application must query it and enforce locally
2	Policy engine (OPA/Cedar) deployed, but policies managed via UI/API
3	Policy-as-Code: policies written in Rego/Cedar, version-controlled in Git, deployed via CI/CD
4	Sidecar enforcement: every microservice delegates authz to a local OPA sidecar via mTLS
5	Perfect decoupling: formally verified policy execution, zero hardcoded authz logic, mathematical proof of policy completeness

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-4.5.1

4.5.1 Data Inputs

Normative source requirement

Does the policy engine have access to real-time, dynamic attributes for decision-making?

Score	Criteria
0	Policy engine only receives user ID and requested action
1	Basic environment context (time of day) included
2	Device posture (EDR status, OS version) fed into policy engine
3	Real-time threat intelligence (e.g., impossible travel anomalies) influences decisions
4	Resource-state awareness: policy evaluates the state of the target resource (e.g., "is this document already classified as restricted?") before authorizing
5	Perfect context: formally verified data pipelines, zero stale attributes, mathematical proof of decision accuracy

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-4.6.1

4.6.1 Compliance Evidence

Normative source requirement

Can the organization prove exactly why an access decision was made at any point in the past?

Score	Criteria
0	Logs only record "Access Granted/Denied"
1	Logs record the user and resource
2	Logs record the RBAC role used
3	ABAC decision logs: records the exact policy version, input attributes, and evaluation path
4	Cryptographic signing of authorization decisions, binding them to tamper-evident ledgers
5	Perfect provenance: formally verified audit trails, zero ability to alter decision history, cryptographic proof of compliance

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-5.1.1

5.1.1 MPAS-PKI

Normative source requirement

• CA Compromise Simulation: 100+ tests attempting to extract the CA private key from the HSM/TPM, verifying the hardware boundary holds.

• Rotation Stress Test: 50+ tests forcing mass certificate expiration across 10,000 microservices, verifying ACME/SPIFFE rotates them without dropping connections.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-5.1.2

5.1.2 MPAS-ABAC

Normative source requirement

• Context Shift Test: 100+ tests changing a user's device posture (e.g., disabling EDR) mid-session, verifying the ABAC engine instantly revokes access on the next request.

• Policy Bypass Test: 50+ tests attempting to bypass the OPA sidecar by calling the microservice directly, verifying mTLS enforcement blocks the unauthenticated request.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results. Context manifests, retrieval traces, source citations, freshness records, conflict decisions, and memory provenance.
ASSESSMENT METHOD	MATURITY SIGNAL
Trace the decision path from identity and policy through execution, attempt bypass and privilege-escalation scenarios, verify revocation behavior, and inspect the resulting evidence record.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-6.1.1

6.1.1 Certificate Authority Compromise

Normative source requirement

Severity: Critical Description: An attacker gains access to the CA's private key. They can now mint valid certificates for any service, establishing persistent, undetectable access to the entire zero-trust network.

Required Controls: 1. CA private keys MUST be generated and stored in FIPS 140-2 Level 3+ HSMs. 2. Keys MUST be non-exportable.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Trace the decision path from identity and policy through execution, attempt bypass and privilege-escalation scenarios, verify revocation behavior, and inspect the resulting evidence record.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-6.1.2

6.1.2 Role Explosion (RBAC Privilege Creep)

Normative source requirement

Severity: High Description: Over time, an accumulates dozens of roles to perform specific tasks. The roles grant broad access. An attacker compromising the user account gains excessive lateral movement capabilities.

Required Controls: 1. Abolish RBAC in favor of ABAC. 2. Evaluate access dynamically based on specific attributes, not broad roles.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Trace the decision path from identity and policy through execution, attempt bypass and privilege-escalation scenarios, verify revocation behavior, and inspect the resulting evidence record.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-6.1.3

6.1.3 Stale Policy Enforcement

Normative source requirement

Severity: High Description: A developer updates the OPA policy in Git, but the CI/CD pipeline fails to deploy it to the sidecars. The microservice enforces outdated, vulnerable authorization logic.

Required Controls: 1. Policy engines MUST report their running policy version/hash. 2. Control plane MUST detect version drift and quarantine non-compliant services.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Architecture records, implemented configuration, test evidence, operational telemetry, exception decisions, and accountable approval.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-ID-0002-6.1.4

6.1.4 Certificate Expiry Outage

Normative source requirement

Severity: High Description: An internal CA issues 1-year certificates. A critical API gateway's certificate expires over the weekend because no one monitored it. All API traffic fails.

Required Controls: 1. Mandatory ACME automation for all certificates. 2. Short certificate lifespans (< 90 days) to force automation.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Architecture records, implemented configuration, test evidence, operational telemetry, exception decisions, and accountable approval.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

Current authority and freshness register

These sources inform interpretation but do not convert this candidate publication into certification, legal compliance, or implementation evidence. Rolling sources must be version-pinned at adoption time.

AUTHORITY	VERSION / FRESHNESS	APPLICABILITY LIMIT
NIST - Digital Identity Guidelines	SP 800-63-4 final, July 2025 Expires 2027-01-24	Federal guidance must be tailored for non-federal risk, jurisdiction, usability, privacy, and operational context.
W3C - Verifiable Credentials Data Model v2.0	W3C Recommendation, May 15 2025 Expires 2027-01-24	Data model conformance does not establish issuer trust, credential truth, or ecosystem governance.
W3C - Decentralized Identifiers (DIDs) v1.0	W3C Recommendation; DID 1.1 remains a Candidate Recommendation Expires 2027-01-24	DID method security, governance, resolution, and privacy properties vary materially.
SPIFFE - SPIFFE Specifications and Workload API	Rolling specification snapshot; SPIRE 1.15.2 documentation Expires 2026-10-24	Implementation and deployment assurance depend on attestation plugins, trust-domain design, and operational controls.

Source lineage appendix

The following text preserves the substantive source draft in a normalized public form. Where it conflicts with the permanent publication identity, candidate status, current authority register, or evidence requirements above, the controlled publication layer governs.

0. Executive Mandate

The architecture of authorization and key management has failed.

Organizations attempt to manage access control using static Role-Based Access Control (RBAC). As the organization scales, RBAC suffers from "role explosion"-thousands of highly specific roles that no one understands, granting excessive standing privileges. Simultaneously, they manage X.509 certificates manually, using Excel spreadsheets to track expiration dates. When a certificate expires, a critical microservice or API gateway goes offline, causing catastrophic outages.

This standard exists because:

1. RBAC is Too Rigid for Dynamic Infrastructure: A user's or agent's authority should not be defined solely by a static group membership. Authority MUST be dynamic, evaluated in real-time based on attributes (identity, device posture, time of day, geolocation, request payload). Systems MUST transition to Attribute-Based Access Control (ABAC).

2. Manual PKI is Operational Suicide: A certificate lifecycle that requires a human to generate a CSR, wait for an email, and manually install the certificate guarantees eventual outages. PKI MUST be fully automated using the ACME (Automated Certificate Management Environment) protocol or SPIFFE, with certificates lasting hours or days, not years.

3. Policy is Code, Not Configuration: Authorization logic hardcoded into application controllers or buried in proprietary IAM UIs is unmanageable. Policies MUST be decoupled from the application, written as code (e.g., Rego/Cedar), version-controlled in Git, and enforced by a centralized policy engine.

4. Private Keys Must Be Hardware-Bound: A certificate authority whose private key exists on a standard virtual machine is a compromised CA waiting to happen. The Root CA and Intermediate CA keys MUST be generated and stored inside Hardware Security Modules (HSMs) or Trusted Platform Modules (TPMs), non-exportable.

This document establishes the Mythos PKI & ABAC Standard (MPAS), a comprehensive, evidence-based methodology for evaluating key management and authorization architectures against automation rigor, dynamic policy evaluation, and cryptographic non-extractability.

Part I. Scope and Applicability

1.1 In Scope

This standard covers the evaluation of:

- Public Key Infrastructure (PKI) and Certificate Authority (CA) hierarchies
- Certificate lifecycle automation (ACME, cert-manager, SPIFFE SVIDs)
- Hardware Security Modules (HSM) and TPM integration
- Attribute-Based Access Control (ABAC) and Policy-as-Code (Open Policy Agent, Cedar)
- The transition from Role-Based Access Control (RBAC) to ABAC
- mTLS (Mutual TLS) enforcement and short-lived identity documents

1.2 Out of Scope

- Human authentication mechanisms like Passkeys/FIDO2 (covered in MYTHOS-ID-0001)
- Network-level Zero Trust architecture (covered in MYTHOS-NET-0004)
- Post-Quantum Cryptography algorithms (covered in MYTHOS-SEC-0001)

1.3 Audience

- Security architects and PKI engineers
- Platform engineers building zero-trust service meshes
- Application architects decoupling authorization logic from microservices
- Compliance teams managing access governance
- Standards bodies seeking a reference PKI and ABAC methodology

Part II. Definitions and Taxonomy

2.1 PKI & Authorization Dimensions

Dimension	Definition
RBAC (Role-Based Access Control)	An access control model where permissions are assigned to roles, and users are assigned to roles. Leads to role explosion and excessive standing privilege.
ABAC (Attribute-Based Access Control)	An access control model where authorization decisions are evaluated dynamically based on attributes of the user, the resource, the action, and the environment (e.g., time, location).
Policy-as-Code	The practice of writing authorization rules in a domain-specific language (e.g., Rego, Cedar), version-controlling them in Git, and enforcing them via a decoupled policy engine.
ACME (Automated Certificate Management Environment)	A protocol (RFC 8555) for automating the validation, issuance, and renewal of X.509 certificates between a client and a Certificate Authority.
HSM (Hardware Security Module)	A physical computing device that safeguards and manages digital keys for strong authentication and provides cryptoprocessing, ensuring private keys are non-exportable.

2.2 The PKI & ABAC Doctrine

> A static role is a standing vulnerability. A manual certificate is an outage. A private key in a file is a breach. If the policy is not code, the authorization is an illusion.

Part III. Evaluation Methodology

3.1 Scoring Model

Each capability is scored from 0 to 5.

Score	Meaning
0	Fails basic task; manual PKI, RBAC only, hardcoded auth logic, software CAs
1	Basic RBAC with automated cert renewal, but no ABAC or HSMs
2	Functional ACME and PKI, but policy logic remains in application code
3	Solid production performance; ABAC via OPA, ACME automated, HSM-backed CAs
4	Strong performance; ephemeral mTLS, context-aware ABAC, GitOps policy deployment
5	Best-in-class; sets the standard for mathematically verified, zero-standing authorization

Weighting

Category	Weight	Rationale
Certificate Lifecycle Automation	20%	Manual cert management causes cascading outages
Hardware-Backed Key Security	15%	Software-based CAs are trivially compromised
RBAC to ABAC Transition	20%	RBAC cannot handle dynamic, zero-trust environments
Policy-as-Code Decoupling	20%	Hardcoded authorization logic is unscalable and insecure
Contextual Telemetry & Enforcement	15%	ABAC requires real-time environmental data (posture, time, location)
Audit & Decision Provenance	10%	Every authorization decision must be cryptographically provable

Total: 100%

A system MUST score at least 3.0 in Certificate Automation and ABAC/Policy-as-Code to be considered for production use.

Part IV. Evaluation Categories

4.1 Certificate Lifecycle Automation (Weight: 20%)

4.1.1 PKI Operations

How are X.509 certificates issued, renewed, and revoked?

Score	Criteria
0	Manual CSRs; certificates last years; tracked in spreadsheets
1	Internal CA exists, but clients must manually request and install certs
2	cert-manager used in Kubernetes, but other infrastructure is manual
3	Full ACME integration across all workloads (on-prem, cloud, edge); certificates last < 90 days
4	Ephemeral identity (SPIFFE SVIDs) with cert TTLs < 1 hour; zero manual rotation
5	Perfect automation: formally verified lifecycle bounds, zero certificate-induced outages, mathematical proof of continuous validity

4.2 Hardware-Backed Key Security (Weight: 15%)

4.2.1 Root of Trust

Are Certificate Authority private keys protected against exfiltration?

Score	Criteria
0	Root CA private key stored on a standard VM or in plaintext
1	Intermediate CA keys password-protected on disk
2	Cloud-managed HSM (AWS CloudHSM, Azure Key Vault Premium) used for CAs
3	Dedicated, on-premises FIPS 140-2 Level 3+ HSMs; keys are non-exportable
4	TPM 2.0 attestation required for all leaf certificates; keys generated in silicon
5	Perfect security: formally verified hardware boundaries, zero ability to extract CA keys, cryptographic proof of key provenance

4.3 RBAC to ABAC Transition (Weight: 20%)

4.3.1 Authorization Model

Are access decisions static (based on groups) or dynamic (based on attributes)?

Score	Criteria
0	Pure RBAC; users assigned to groups ("Admins", "Users") with broad permissions
1	RBAC with some coarse context (e.g., IP-based restrictions)
2	Hybrid model: RBAC for broad access, ABAC for sensitive resources
3	Pure ABAC: roles abolished; access evaluated based on user attributes, resource tags, and environment
4	Real-time attribute evaluation: access changes dynamically if device posture degrades or geolocation shifts

Score	Criteria
5	Perfect model: formally verified policy bounds, zero standing privilege, mathematical proof of least-privilege enforcement

4.4 Policy-as-Code Decoupling (Weight: 20%)

4.4.1 Enforcement Architecture

Is authorization logic decoupled from the application codebase?

Score	Criteria
0	Authorization logic hardcoded in application controllers (e.g., `if user.role == 'admin'`)
1	Centralized IAM, but application must query it and enforce locally
2	Policy engine (OPA/Cedar) deployed, but policies managed via UI/API
3	Policy-as-Code: policies written in Rego/Cedar, version-controlled in Git, deployed via CI/CD
4	Sidecar enforcement: every microservice delegates authz to a local OPA sidecar via mTLS
5	Perfect decoupling: formally verified policy execution, zero hardcoded authz logic, mathematical proof of policy completeness

4.5 Contextual Telemetry and Enforcement (Weight: 15%)

4.5.1 Data Inputs

Does the policy engine have access to real-time, dynamic attributes for decision-making?

Score	Criteria
0	Policy engine only receives user ID and requested action
1	Basic environment context (time of day) included
2	Device posture (EDR status, OS version) fed into policy engine
3	Real-time threat intelligence (e.g., impossible travel anomalies) influences decisions
4	Resource-state awareness: policy evaluates the state of the target resource (e.g., "is this document already classified as restricted?") before authorizing
5	Perfect context: formally verified data pipelines, zero stale attributes, mathematical proof of decision accuracy

4.6 Audit and Decision Provenance (Weight: 10%)

4.6.1 Compliance Evidence

Can the organization prove exactly why an access decision was made at any point in the past?

Score	Criteria
0	Logs only record "Access Granted/Denied"
1	Logs record the user and resource
2	Logs record the RBAC role used
3	ABAC decision logs: records the exact policy version, input attributes, and evaluation path
4	Cryptographic signing of authorization decisions, binding them to tamper-evident ledgers
5	Perfect provenance: formally verified audit trails, zero ability to alter decision history, cryptographic proof of compliance

Part V. The Mythos PKI & ABAC Suite (MPAS)

Traditional IAM benchmarks measure authentication latency. The MPAS evaluates certificate rotation resilience, policy engine performance, and ABAC decision accuracy.

5.1 Suite Composition

5.1.1 MPAS-PKI

• CA Compromise Simulation: 100+ tests attempting to extract the CA private key from the HSM/TPM, verifying the hardware boundary holds.

• Rotation Stress Test: 50+ tests forcing mass certificate expiration across 10,000 microservices, verifying ACME/SPIFFE rotates them without dropping connections.

5.1.2 MPAS-ABAC

• Context Shift Test: 100+ tests changing a user's device posture (e.g., disabling EDR) mid-session, verifying the ABAC engine instantly revokes access on the next request.

• Policy Bypass Test: 50+ tests attempting to bypass the OPA sidecar by calling the microservice directly, verifying mTLS enforcement blocks the unauthenticated request.

5.2 Execution Protocol

```
1. Deploy PKI infrastructure and OPA/Cedar policy engine
2. Execute complex workflows requiring dynamic attribute evaluation
3. Run MPAS suite (automated CA attack + context shift injection)
4. Score each category 0-5
5. Check for manual certificate tracking or hardcoded RBAC logic (instant disqualification)
6. If all thresholds met: approve for production
```

Part VI. Security Threat Model for PKI & ABAC

6.1 Threat Catalog

6.1.1 Certificate Authority Compromise

Severity: Critical Description: An attacker gains access to the CA's private key. They can now mint valid certificates for any service, establishing persistent, undetectable access to the entire zero-trust network.

Required Controls:

1. CA private keys MUST be generated and stored in FIPS 140-2 Level 3+ HSMs.

2. Keys MUST be non-exportable.

6.1.2 Role Explosion (RBAC Privilege Creep)

Severity: High Description: Over time, an accumulates dozens of roles to perform specific tasks. The roles grant broad access. An attacker compromising the user account gains excessive lateral movement capabilities.

Required Controls:

1. Abolish RBAC in favor of ABAC.

2. Evaluate access dynamically based on specific attributes, not broad roles.

6.1.3 Stale Policy Enforcement

Severity: High Description: A developer updates the OPA policy in Git, but the CI/CD pipeline fails to deploy it to the sidecars. The microservice enforces outdated, vulnerable authorization logic.

Required Controls:

1. Policy engines MUST report their running policy version/hash.

2. Control plane MUST detect version drift and quarantine non-compliant services.

6.1.4 Certificate Expiry Outage

Severity: High Description: An internal CA issues 1-year certificates. A critical API gateway's certificate expires over the weekend because no one monitored it. All API traffic fails.

Required Controls:

1. Mandatory ACME automation for all certificates.

2. Short certificate lifespans (< 90 days) to force automation.

Part VII. The Mythos PKI & ABAC Standard

7.1 The Mythos PKI & ABAC Doctrine

```
A static role is a standing vulnerability.
A manual certificate is an outage.

A private key in a file is a breach.
If the policy is not code, the authorization is an illusion.

Access may be requested.
Dynamic authorization must be absolute.
```

7.2 Selection Rules

1. No PKI or ABAC architecture enters production without passing MPAS.
2. No architecture is approved if it relies on manual certificate management.
3. No CA is approved if its private keys are not hardware-bound (HSM/TPM).
4. No architecture is approved if it relies solely on static RBAC roles.
5. No architecture is approved if authorization logic is hardcoded in the application.

7.3 The Prime Directive for PKI & ABAC Architecture

> Automate the ACME, do not track the spreadsheet. > Bind the key to the silicon, do not store the PEM. > Evaluate the attribute, do not trust the role. > Decouple the policy, do not hardcode the logic.

```
Academic benchmarks measure OPA evaluation latency.
Applied benchmarks measure certificate rotation resilience.
Security benchmarks measure HSM non-extractability.
Operational benchmarks measure ABAC context shift response.

> Certificates may be ephemeral.
> Authorization must be absolute.
```

End of Standard MYTHOS-ID-0002

End of controlled publication

MYTHOS-ID-0002 remains a Candidate Standard until technical review, security and privacy review, accessibility review, current-source validation, and formal approval are recorded.

ENGINEERING STANDARDS LIBRARY / LEARNING AND WORKFORCE

Adaptive Learning, Skill Graphs & Cyber Lab Standard

The architecture of technical education and capability tracking has failed.

PERMANENT DOCUMENT ID

MYTHOS-EDU-0001

PUBLICATION

Candidate Standard
Version 1.0.0-candidate

ASSURANCE PROFILE

Standard / Advanced
High Assurance

PERMANENT PUBLICATION IDENTITY

Adaptive Learning, Skill Graphs & Cyber Lab Standard

DOCUMENT ID
MYTHOS-EDU-0001

VERSION
1.0.0-candidate

STATUS
Candidate Standard

PUBLISHER
Mythos Systems

PUBLICATION DATE
2026-07-24

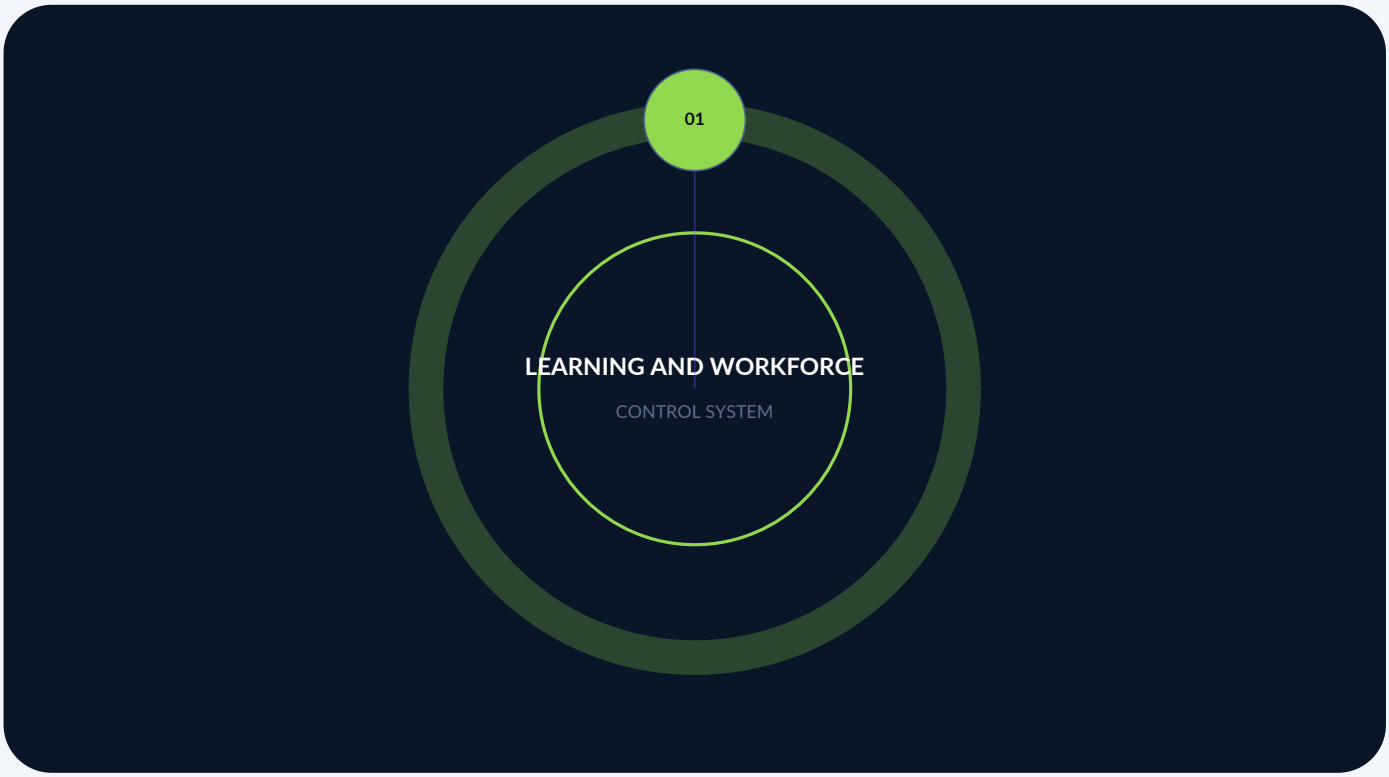
SCHEDULED REVIEW
2027-07-24

12	6	4	1
ATOMIC CRITERIA	ASSURANCE VIEWS	IMPLEMENTATION PROFILES	PERMANENT IDENTITY

Publication doctrine. This standard is a permanent library identity. Interpretation must preserve its recorded evidence, controlled exceptions, formal revision history, and explicit relationship to companion standards.

DOMAIN CONTROL SYSTEM

MYTHOS-EDU-0001 inside the learning and workforce constellation

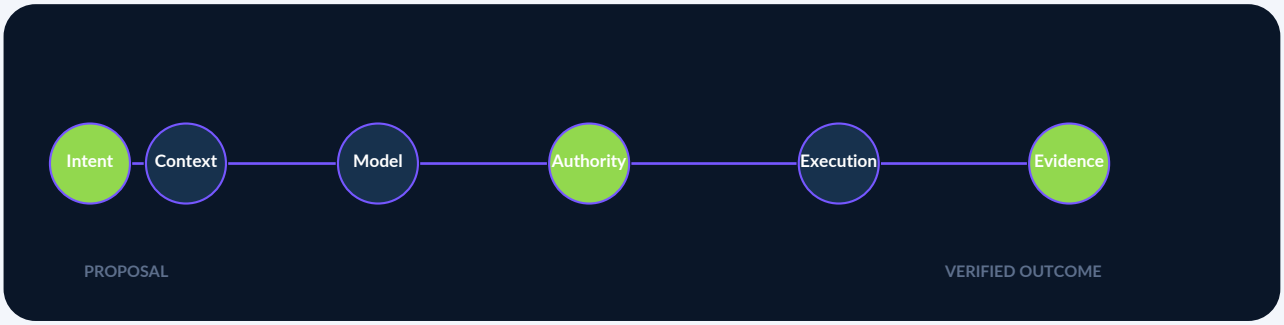


Connected assurance	Architecture, implementation, operations, evidence, and lifecycle are evaluated as one controlled system.
Specialization rule	Companion standards may add precision, but cannot silently weaken this publication.

Executive orientation

Defines adaptive learning, skill graphs, competency evidence, cyber labs, assessment, accessibility, credential portability, safety, and workforce-assurance requirements.

WHY THIS STANDARD EXISTS	The architecture of technical education and capability tracking has failed.
OPERATIONAL SCOPE	This standard covers the evaluation of: - AI-driven adaptive learning engines and dynamic curriculum generation - Semantic Skill Graphs and competency ontologies - Ephemeral cyber range architectures (containerized, microVM) - Evidence-based assessment (automated grading of live infrastructural changes) - Zero-trust isolation for training environments - Integration of operational telemetry (e.g., SIEM, Git) into skill verification
PRIMARY AUDIENCE	- Security engineering leads and capability managers - Platform engineers building cyber ranges and ephemeral lab infrastructure - AI engineers building adaptive tutoring systems - Learning and Development (L&D) directors modernizing technical training - Standards bodies seeking a reference education methodology



Assurance principle. Model capability, data quality, identity, authority, operational evidence, and human accountability are treated as a connected system. A high aggregate score cannot compensate for a failed mandatory safety or authority boundary.

Contents

Evaluation criterion index

This standard contains 12 atomic criteria. Each criterion is independently testable and requires retained evidence.

ID	EVALUATION CRITERION
4.1.1	Personalized Learning
4.2.1	Lab Provisioning
4.3.1	Competency Verification
4.4.1	Capability Mapping
4.5.1	Real-World Relevance
4.6.1	Boundary Enforcement
5.1.1	MEST-Lab
5.1.2	MEST-Adaptation
6.1.1	Lab-to-Production Pivot
6.1.2	Assessment Fraud (The Fake Skill)
6.1.3	Stale Curriculum (The Obsolete Defense)
6.1.4	Graph Spoofing (HR System Hack)

MYTHOS-EDU-0001-4.1.1

4.1.1 Personalized Learning

Normative source requirement

Is the curriculum static, or does it adapt to the learner?

Score	Criteria
0	One-size-fits-all video modules and PDFs
1	Branching logic based on quiz scores (e.g., if < 80%, review module)
2	Basic AI recommendations for next courses
3	Real-time adaptive engine: AI analyzes learner's lab performance and dynamically generates new scenarios targeting specific weaknesses
4	Dynamic content generation: RAG over live threat intelligence (CISA KEV) creates zero-day scenarios not pre-built by instructors
5	Perfect adaptation: formally verified cognitive load modeling, AI-driven curriculum generation, mathematical proof of optimal learning paths

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Context manifests, retrieval traces, source citations, freshness records, conflict decisions, and memory provenance.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-4.2.1

4.2.1 Lab Provisioning

Normative source requirement

How are hands-on training environments deployed?

Score	Criteria
0	Persistent VMs shared among students; manual resets required
1	Dedicated VMs per student, but take 5+ minutes to boot
2	Containerized labs, but lack complex network topologies (e.g., BGP, firewalls)
3	Ephemeral microVM/container hybrid ranges spun up in < 10 seconds; completely destroyed on session end
4	Infrastructure-as-Code (IaC) based ranges mirroring production environments (e.g., a simulated the governed reference platform agent runtime)
5	Perfect architecture: formally verified ephemeral bounds, zero drift between sessions, mathematical proof of topological fidelity

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Model and dataset cards, lineage, licenses, evaluation reports, release approvals, and rollback artifacts.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-4.3.1

4.3.1 Competency Verification

Normative source requirement

How is a learner's skill measured?

Score	Criteria
0	Multiple-choice quizzes at the end of a module
1	Fill-in-the-blank CLI commands
2	Instructor manually reviews the state of the lab VM
3	Automated grading: system checks the live state of the lab (e.g., API verifies RPKI is configured and BGP session is up)
4	Behavioral assessment: system evaluates the process (e.g., checking command history to ensure safe practices were used)
5	Perfect verification: formally verified assessment bounds, zero ability to cheat, cryptographic attestation of competency

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Policy configuration, identity or trust records, authorization decisions, revocation evidence, and adversarial test results.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-4.4.1

4.4.1 Capability Mapping

Normative source requirement

Are skills tracked as a flat list or a semantic graph?

Score	Criteria
0	Flat spreadsheet of skills ("Tier 1 Analyst", "Knows Python")
1	Tagging system in an LMS, but unstructured
2	Basic hierarchical taxonomy of skills
3	Semantic Skill Graph: competencies mapped as nodes, linked to specific tools, tasks, and verified evidence edges
4	Automated graph updates: passing a lab automatically updates the graph and recommends new operational tasks
5	Perfect mapping: formally verified ontology, zero gap between graph state and actual operational capability, real-time skill parity with production needs

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Trace schemas, immutable event records, integrity verification, retention policy, and operator queries.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-4.5.1

4.5.1 Real-World Relevance

Normative source requirement

Does the training reflect the actual production architecture and threat landscape?

Score	Criteria
0	Generic cyber ranges (e.g., Metasploitable) disconnected from corporate reality
1	Custom scenarios, but updated annually
2	Scenarios mapped to MITRE ATT&CK, but lack production tooling
3	Training environments replicate the enterprise architecture (e.g., using internal Terraform modules and the governed reference platform agents)
4	AI automatically ingests the organization's actual SIEM alerts and generates training scenarios based on real internal incidents
5	Perfect alignment: formally verified threat parity, zero gap between training and operational reality, mathematical proof of readiness

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records. Model and dataset cards, lineage, licenses, evaluation reports, release approvals, and rollback artifacts.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-4.6.1

4.6.1 Boundary Enforcement

Normative source requirement

Can a learner's exploit escape the cyber range?

Score	Criteria
0	Labs on the same VLAN as corporate network
1	Labs on a dedicated VLAN, but rely on software firewalls
2	Labs in an isolated cloud VPC, but allow outbound internet
3	Strict air-gapped lab networks; zero-trust microsegmentation per learner
4	Egress firewalls block all outbound traffic; malware analysis requires a local internet simulator
5	Perfect isolation: formally verified zero-trust boundaries, zero ability for lab exploits to reach corporate networks, cryptographic attestation of session containment

---

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records.
ASSESSMENT METHOD	MATURITY SIGNAL
Run a version-pinned workload with representative inputs, record distributions rather than a single score, repeat under failure and load, and compare reproduced results with the stated acceptance threshold.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-5.1.1

5.1.1 MEST-Lab

Normative source requirement

• Provisioning Stress Test: 100+ tests spinning up 1,000 concurrent ephemeral cyber ranges, verifying they boot in < 10 seconds and are strictly isolated from one another.

• Grading Accuracy Test: 50+ tests executing complex configurations (e.g., implementing zero-trust mTLS), verifying the automated grading engine correctly detects both successes and subtle misconfigurations.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Reproducible benchmark results, workload definitions, raw outputs, and variance records.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-5.1.2

5.1.2 MEST-Adaptation

Normative source requirement

• Dynamic Generation Test: 100+ tests feeding a novel CVE to the AI engine, verifying it dynamically generates a working lab environment and assessment criteria on the fly.

• Isolation Break Test: 50+ tests attempting to pivot from a compromised learner VM to the host infrastructure or another learner's VM, verifying the microVM/container boundary blocks the attack.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Skill graph, assessment blueprint, learner evidence, lab isolation record, accessibility results, and credential metadata.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-6.1.1

6.1.1 Lab-to-Production Pivot

Normative source requirement

Severity: Critical Description: A learner is practicing advanced exploitation techniques in a cyber range. Due to a misconfigured VLAN or lack of microsegmentation, the learner (or a malicious actor in the training network) pivots from the lab into the corporate production network, deploying ransomware.

Required Controls: 1. Labs MUST run in ephemeral microVMs (e.g., Firecracker) with strict network isolation. 2. Lab networks MUST be physically or cryptographically air-gapped from production.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Model and dataset cards, lineage, licenses, evaluation reports, release approvals, and rollback artifacts.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-6.1.2

6.1.2 Assessment Fraud (The Fake Skill)

Normative source requirement

Severity: High Description: An employee clicks "Next" through a 50-slide presentation and passes a multiple-choice quiz. Their HR profile is updated to "Certified Kubernetes Administrator." They are later assigned to a production incident and cause an outage due to lack of actual competence.

Required Controls: 1. Abolish multiple-choice quizzes for technical roles. 2. Assessment MUST be evidence-based, verifying live system state via API.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Trace schemas, immutable event records, integrity verification, retention policy, and operator queries. Skill graph, assessment blueprint, learner evidence, lab isolation record, accessibility results, and credential metadata.
ASSESSMENT METHOD	MATURITY SIGNAL
Exercise crash, retry, duplicate, reorder, partition, and restore scenarios; compare authoritative state before and after replay; confirm idempotency and recovery evidence.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-6.1.3

6.1.3 Stale Curriculum (The Obsolete Defense)

Normative source requirement

Severity: High Description: The organization's training covers defending against Log4j, but a new critical zero-day (e.g., xz-utils backdoor) emerges. The training will not be updated for 6 months, leaving the workforce unprepared.

Required Controls: 1. AI-driven dynamic curriculum generation utilizing RAG over live CISA KEV and internal threat intel.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Context manifests, retrieval traces, source citations, freshness records, conflict decisions, and memory provenance. Model and dataset cards, lineage, licenses, evaluation reports, release approvals, and rollback artifacts.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

MYTHOS-EDU-0001-6.1.4

6.1.4 Graph Spoofing (HR System Hack)

Normative source requirement

Severity: Medium Description An attacker compromises the LMS database and manually alters their Skill Graph to grant themselves "Tier 3 SOC Analyst" privileges, leading to a role escalation.

Required Controls: 1. Skill Graph updates MUST be cryptographically signed by the ephemeral lab's grading engine. 2. Manual overrides must require dual-control approval.

CONTROL INTENT	REQUIRED EVIDENCE
Create a measurable, reviewable control that can be verified independently of model or vendor claims.	Skill graph, assessment blueprint, learner evidence, lab isolation record, accessibility results, and credential metadata.
ASSESSMENT METHOD	MATURITY SIGNAL
Inspect the architecture and configured control, execute normal and adverse scenarios, verify measurable outcomes, sample retained evidence, and confirm that exceptions are time-bound and approved.	0 absent | 1 ad hoc | 2 repeatable | 3 controlled | 4 measured | 5 continuously verified

Decision rule. A criterion is not satisfied by documentation alone when the claim concerns runtime behavior. Reproduced evidence, responsible ownership, known limitations, and controlled exceptions are required.

Current authority and freshness register

These sources inform interpretation but do not convert this candidate publication into certification, legal compliance, or implementation evidence. Rolling sources must be version-pinned at adoption time.

AUTHORITY	VERSION / FRESHNESS	APPLICABILITY LIMIT
NIST - Artificial Intelligence Risk Management Framework (AI RMF 1.0)	NIST AI 100-1; current framework, revision underway Expires 2026-10-24	Voluntary risk-management framework; does not establish product certification or implementation effectiveness.
W3C - Verifiable Credentials Data Model v2.0	W3C Recommendation, May 15 2025 Expires 2027-01-24	Data model conformance does not establish issuer trust, credential truth, or ecosystem governance.
1EdTech - Open Badges	Open Badges 3.0 final Expires 2027-01-24	Credential interoperability does not establish assessment validity, issuer quality, or workforce acceptance.
1EdTech - Comprehensive Learner Record Standard	CLR 2.0 final Expires 2027-01-24	Record portability does not establish competency, authenticity of evidence, or equitable assessment.

Source lineage appendix

The following text preserves the substantive source draft in a normalized public form. Where it conflicts with the permanent publication identity, candidate status, current authority register, or evidence requirements above, the controlled publication layer governs.

0. Executive Mandate

The architecture of technical education and capability tracking has failed.

Organizations attempt to train engineers and security operators using static videos and multiple-choice quizzes hosted on legacy Learning Management Systems (LMS). They track competency via a flat spreadsheet of "skills" that are neither verified nor mapped to actual operational tasks. When an engineer is assigned to an incident, their actual capability rarely matches their "certified" status, leading to catastrophic failure under pressure. Furthermore, hands-on cyber training relies on brittle, persistent virtual machines that take minutes to boot and share a flat network, allowing a single learner's mistake to break the environment for the class.

This standard exists because:

1. Compliance Training is Not Competence: Passing a multiple-choice quiz on "How to Mitigate BGP Hijacking" proves the user can read; it does not prove they can configure RPKI on a live router. Systems MUST utilize evidence-based, hands-on cyber labs where assessment is derived from the successful execution of tasks in a live environment.

2. Static Curriculum is Obsolete: Threat landscapes and infrastructure paradigms change weekly. Training content MUST be dynamically generated and updated by AI (RAG over current threat intelligence) and adapted in real-time to the learner's specific cognitive gaps.

3. Skills Must Be a Formal Graph: A flat list of skills (e.g., "Knows Python") is unstructured data. Systems MUST implement a Semantic Skill Graph, mapping competencies to specific tools, architectural patterns, and verifiable operational evidence (e.g., "Proven ability to implement Pydantic AI tool schemas").

4. Cyber Labs Must Be Ephemeral and Isolated: Training environments that persist for days invite cross-contamination and drift. Labs MUST be ephemeral, spun up in seconds via containers/microVMs, strictly isolated per learner, and destroyed upon completion.

This document establishes the Mythos Education Standard (MEST), a comprehensive, evidence-based methodology for evaluating learning architectures against adaptive AI rigor, semantic skill mapping, and zero-trust cyber range resilience.

Part I. Scope and Applicability

1.1 In Scope

This standard covers the evaluation of:

- AI-driven adaptive learning engines and dynamic curriculum generation
- Semantic Skill Graphs and competency ontologies
- Ephemeral cyber range architectures (containerized, microVM)
- Evidence-based assessment (automated grading of live infrastructural changes)
- Zero-trust isolation for training environments
- Integration of operational telemetry (e.g., SIEM, Git) into skill verification

1.2 Out of Scope

- General HR onboarding and compliance tracking (e.g., workplace harassment training)
- General academic university curricula (unless operating sovereign labs)
- General platform engineering (covered in MYTHOS-PLT-0003)

1.3 Audience

- Security engineering leads and capability managers
- Platform engineers building cyber ranges and ephemeral lab infrastructure
- AI engineers building adaptive tutoring systems
- Learning and Development (L&D) directors modernizing technical training
- Standards bodies seeking a reference education methodology

Part II. Definitions and Taxonomy

2.1 Education Dimensions

Dimension	Definition
Adaptive Learning Engine	An AI-driven system that dynamically alters the difficulty, topic, and format of training based on real-time analysis of the learner's performance and cognitive load.
Semantic Skill Graph	A knowledge graph mapping technical competencies (nodes) to operational tasks, tools, and verified evidence (edges), replacing flat spreadsheets.
Ephemeral Cyber Range	A fully isolated, virtualized training environment (utilizing containers or microVMs) that is provisioned instantly for a specific task and destroyed immediately upon completion.
Evidence-Based Assessment	The practice of grading a learner not on quizzes, but by programmatically verifying the state of a live system (e.g., checking if a firewall rule was correctly applied via API).
Dynamic Curriculum Generation	The use of RAG (Retrieval-Augmented Generation) over current threat intelligence and architecture standards to generate training scenarios on the fly.

2.2 The Education Doctrine

> A multiple-choice quiz is a test of reading, not competence. A static PDF is a liability. A persistent training VM is a security risk. If the skill is not a graph mapped to evidence, it is a guess.

Part III. Evaluation Methodology

3.1 Scoring Model

Each capability is scored from 0 to 5.

Score	Meaning
0	Fails basic task; static LMS, multiple-choice, flat skills, persistent VMs
1	Basic hands-on labs, but manual grading and static curriculum
2	Functional cyber range, but lacks adaptive AI or semantic skill mapping
3	Solid production performance; AI adaptive learning, ephemeral ranges, evidence-based grading, semantic skill graph
4	Strong performance; AI-generated scenarios, automated evidence ingestion from production Git/SIEM
5	Best-in-class; sets the standard for mathematically verified, autonomous capability development

Weighting

Category	Weight	Rationale
Adaptive AI & Dynamic Curriculum	20%	Static content is obsolete the day it is published
Ephemeral Cyber Range Architecture	20%	Persistent labs drift, break, and invite cross-contamination
Evidence-Based Assessment	20%	Quizzes cannot verify infrastructural competency
Semantic Skill Graph Integration	15%	Flat skill matrices cannot map to complex operational needs
Operational Alignment & Threat Intel	15%	Training must reflect the actual production environment
Isolation & Security of Labs	10%	A learner's exploit must never reach the corporate network

Total: 100%

A system MUST score at least 3.0 in Ephemeral Cyber Ranges and Evidence-Based Assessment to be considered for production use.

Part IV. Evaluation Categories

4.1 Adaptive AI and Dynamic Curriculum (Weight: 20%)

4.1.1 Personalized Learning

Is the curriculum static, or does it adapt to the learner?

Score	Criteria
0	One-size-fits-all video modules and PDFs
1	Branching logic based on quiz scores (e.g., if < 80%, review module)
2	Basic AI recommendations for next courses
3	Real-time adaptive engine: AI analyzes learner's lab performance and dynamically generates new scenarios targeting specific weaknesses
4	Dynamic content generation: RAG over live threat intelligence (CISA KEV) creates zero-day scenarios not pre-built by instructors
5	Perfect adaptation: formally verified cognitive load modeling, AI-driven curriculum generation, mathematical proof of optimal learning paths

4.2 Ephemeral Cyber Range Architecture (Weight: 20%)

4.2.1 Lab Provisioning

How are hands-on training environments deployed?

Score	Criteria
0	Persistent VMs shared among students; manual resets required
1	Dedicated VMs per student, but take 5+ minutes to boot
2	Containerized labs, but lack complex network topologies (e.g., BGP, firewalls)
3	Ephemeral microVM/container hybrid ranges spun up in < 10 seconds; completely destroyed on session end
4	Infrastructure-as-Code (IaC) based ranges mirroring production environments (e.g., a simulated the governed reference platform agent runtime)
5	Perfect architecture: formally verified ephemeral bounds, zero drift between sessions, mathematical proof of topological fidelity

4.3 Evidence-Based Assessment (Weight: 20%)

4.3.1 Competency Verification

How is a learner's skill measured?

Score	Criteria
0	Multiple-choice quizzes at the end of a module
1	Fill-in-the-blank CLI commands
2	Instructor manually reviews the state of the lab VM
3	Automated grading: system checks the live state of the lab (e.g., API verifies RPKI is configured and BGP session is up)
4	Behavioral assessment: system evaluates the process (e.g., checking command history to ensure safe practices were used)

Score	Criteria
5	Perfect verification: formally verified assessment bounds, zero ability to cheat, cryptographic attestation of competency

4.4 Semantic Skill Graph Integration (Weight: 15%)

4.4.1 Capability Mapping

Are skills tracked as a flat list or a semantic graph?

Score	Criteria
0	Flat spreadsheet of skills ("Tier 1 Analyst", "Knows Python")
1	Tagging system in an LMS, but unstructured
2	Basic hierarchical taxonomy of skills
3	Semantic Skill Graph: competencies mapped as nodes, linked to specific tools, tasks, and verified evidence edges
4	Automated graph updates: passing a lab automatically updates the graph and recommends new operational tasks
5	Perfect mapping: formally verified ontology, zero gap between graph state and actual operational capability, real-time skill parity with production needs

4.5 Operational Alignment and Threat Intel (Weight: 15%)

4.5.1 Real-World Relevance

Does the training reflect the actual production architecture and threat landscape?

Score	Criteria
0	Generic cyber ranges (e.g., Metasploitable) disconnected from corporate reality
1	Custom scenarios, but updated annually
2	Scenarios mapped to MITRE ATT&CK, but lack production tooling
3	Training environments replicate the enterprise architecture (e.g., using internal Terraform modules and the governed reference platform agents)
4	AI automatically ingests the organization's actual SIEM alerts and generates training scenarios based on real internal incidents
5	Perfect alignment: formally verified threat parity, zero gap between training and operational reality, mathematical proof of readiness

4.6 Isolation and Security of Labs (Weight: 10%)

4.6.1 Boundary Enforcement

Can a learner's exploit escape the cyber range?

Score	Criteria
0	Labs on the same VLAN as corporate network
1	Labs on a dedicated VLAN, but rely on software firewalls
2	Labs in an isolated cloud VPC, but allow outbound internet
3	Strict air-gapped lab networks; zero-trust microsegmentation per learner
4	Egress firewalls block all outbound traffic; malware analysis requires a local internet simulator

Score	Criteria
5	Perfect isolation: formally verified zero-trust boundaries, zero ability for lab exploits to reach corporate networks, cryptographic attestation of session containment

Part V. The Mythos Education Suite (MEST)

Traditional LMS benchmarks measure course completion rates. The MEST evaluates ephemeral lab resilience, automated grading accuracy, and semantic graph fidelity.

5.1 Suite Composition

5.1.1 MEST-Lab

• Provisioning Stress Test: 100+ tests spinning up 1,000 concurrent ephemeral cyber ranges, verifying they boot in < 10 seconds and are strictly isolated from one another.

• Grading Accuracy Test: 50+ tests executing complex configurations (e.g., implementing zero-trust mTLS), verifying the automated grading engine correctly detects both successes and subtle misconfigurations.

5.1.2 MEST-Adaptation

• Dynamic Generation Test: 100+ tests feeding a novel CVE to the AI engine, verifying it dynamically generates a working lab environment and assessment criteria on the fly.

• Isolation Break Test: 50+ tests attempting to pivot from a compromised learner VM to the host infrastructure or another learner's VM, verifying the microVM/container boundary blocks the attack.

5.2 Execution Protocol

```
1. Deploy adaptive learning architecture with ephemeral ranges and skill graph
2. Assign a complex, multi-stage architectural task to a simulated learner
3. Run MEST suite (automated provisioning + dynamic generation + isolation tests)
4. Score each category 0-5
5. Check for multiple-choice assessments or persistent VMs (instant disqualification)
6. If all thresholds met: approve for production
```

Part VI. Security Threat Model for Education Systems

6.1 Threat Catalog

6.1.1 Lab-to-Production Pivot

Severity: Critical Description: A learner is practicing advanced exploitation techniques in a cyber range. Due to a misconfigured VLAN or lack of microsegmentation, the learner (or a malicious actor in the training network) pivots from the lab into the corporate production network, deploying ransomware.

Required Controls:

1. Labs MUST run in ephemeral microVMs (e.g., Firecracker) with strict network isolation.

2. Lab networks MUST be physically or cryptographically air-gapped from production.

6.1.2 Assessment Fraud (The Fake Skill)

Severity: High Description: An employee clicks "Next" through a 50-slide presentation and passes a multiple-choice quiz. Their HR profile is updated to "Certified Kubernetes Administrator." They are later assigned to a production incident and cause an outage due to lack of actual competence.

Required Controls:

1. Abolish multiple-choice quizzes for technical roles.

2. Assessment MUST be evidence-based, verifying live system state via API.

6.1.3 Stale Curriculum (The Obsolete Defense)

Severity: High Description: The organization's training covers defending against Log4j, but a new critical zero-day (e.g., xz-utils backdoor) emerges. The training will not be updated for 6 months, leaving the workforce unprepared.

Required Controls:

1. AI-driven dynamic curriculum generation utilizing RAG over live CISA KEV and internal threat intel.

6.1.4 Graph Spoofing (HR System Hack)

Severity: Medium Description An attacker compromises the LMS database and manually alters their Skill Graph to grant themselves "Tier 3 SOC Analyst" privileges, leading to a role escalation.

Required Controls:

1. Skill Graph updates MUST be cryptographically signed by the ephemeral lab's grading engine.

2. Manual overrides must require dual-control approval.

Part VII. The Mythos Education Standard

7.1 The Mythos Education Doctrine

```
A multiple-choice quiz is a test of reading, not competence.
A static PDF is a liability.

A persistent training VM is a security risk.
If the skill is not a graph mapped to evidence, it is a guess.

Training may be educational.
Operational readiness must be absolute.
```

7.2 Selection Rules

1. No education architecture enters production without passing MEST.

2. No architecture is approved if it relies on multiple-choice assessments for technical roles.

3. No architecture is approved if its cyber ranges are persistent and shared.

4. No architecture is approved without AI-driven dynamic curriculum generation.

5. No architecture is approved if skills are tracked as a flat list rather than a semantic graph.

7.3 The Prime Directive for Education Architecture

> Adapt the curriculum, do not static the PDF. > Destroy the VM, do not persist the state. > Verify the config, do not quiz the theory. > Graph the evidence, do not guess the competence.

```
Academic benchmarks measure course completion rates.
Applied benchmarks measure ephemeral lab isolation.
Security benchmarks measure evidence-based grading rigor.
Operational benchmarks measure semantic skill graph fidelity.

> Learning may be continuous.
> Competence must be absolute.
```

End of Standard MYTHOS-EDU-0001

End of controlled publication

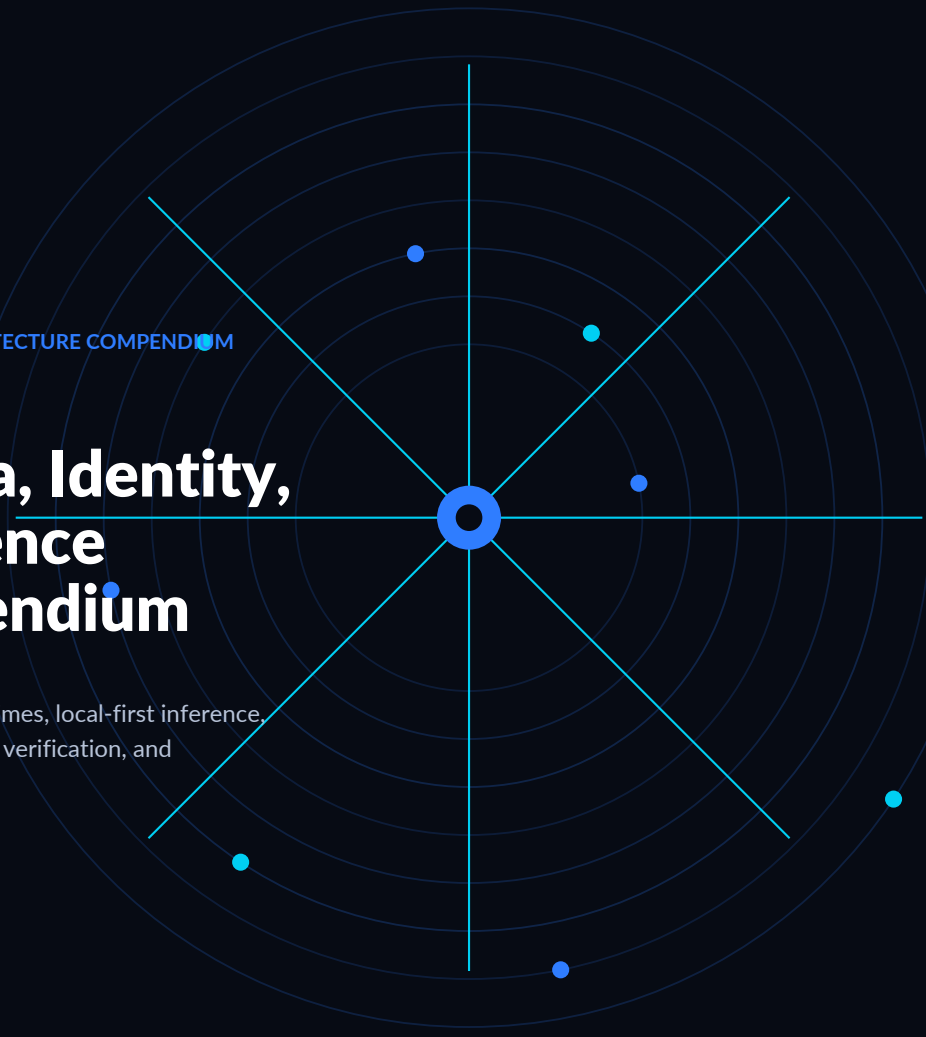
MYTHOS-EDU-0001 remains a Candidate Standard until technical review, security and privacy review, accessibility review, current-source validation, and formal approval are recorded.

M MYTHOS SYSTEMS

ENGINEERING STANDARDS LIBRARY / REFERENCE ARCHITECTURE COMPENDIUM

AI, Knowledge, Data, Identity, and Learning Reference Architecture Compendium

Candidate architectures for governed multi-agent runtimes, local-first inference, and AI-assisted engineering with identity, policy, state, verification, and evidence.



PERMANENT DOCUMENT ID
MYTHOS-RA-COMP-AIKDIL-0001

PUBLICATION
Reference Architecture Compendium

STATUS
Candidate Architecture Compendium

Contents

Architecture doctrine

Reference architectures show coherent implementation patterns; they do not replace workload-specific design, threat modeling, capacity analysis, privacy review, or proof-of-concept evidence.

Architecture register

ARCHITECTURE	ROLE IN THIS LIBRARY
RA-001 - RA-001: Governed Multi-Agent Runtime (the governed reference platform Architecture)	Governs multi-agent runtime boundaries, authority separation, mission durability, verification, and evidence.
RA-006 - RA-006: Local-First AI Inference Cluster (Edge/Node design)	Defines local-first inference clusters, sovereign edge operation, model serving, and controlled acceleration.
RA-007 - RA-007: AI-Assisted Software Engineering Pipeline (Cursor/IDE integration)	Connects model-assisted engineering to repository controls, testing, security review, release evidence, and human approval.

Integrated architecture layers

ARCHITECTURE LAYER	CORE COMPONENTS	TRUST BOUNDARY
Experience	Desktop, CLI, API, voice, vision, and approval interfaces.	Untrusted input is separated from authenticated user intent and authority.
Mission and planning	Typed objectives, decomposition, DAG, budgets, checkpoints, and supervisors.	Plans remain proposals until policy and approval conditions are satisfied.
Knowledge and memory	Context compiler, retrieval, provenance, working state, episodic and semantic memory.	Source authority, tenant scope, sensitivity, and deletion are enforced.
Model fabric	Local and hosted serving, routing, profiles, adaptation, and evaluation.	Models cannot access secrets or execution capabilities except through controlled services.
Identity and policy	Human and workload identity, attestation, capability grants, secrets, and approvals.	Authorization is independent, deny-capable, revocable, and evidence-producing.
Execution	Sandboxed tools, code runners, processes, repositories, workflows, and transactional services.	Blast radius, network, filesystem, credentials, and side effects are constrained.
Verification	Tests, simulations, policy checks, evaluators, observed state, and human review.	Verification methods remain sufficiently independent of the proposing path.
Evidence and operations	Event store, traces, metrics, audit, incidents, SLOs, cost, and recovery.	Evidence integrity and operational access are protected from the model runtime.

End-to-end controlled sequence

- 1. A human or trusted system submits a typed objective, constraints, consequence class, and acceptance criteria.
- 2. Context and knowledge services assemble an authority-ranked, provenance-preserving packet.
- 3. A planner proposes a mission graph without receiving execution authority.

- 4. Identity and policy services authenticate the principal and issue scoped, expiring capability grants.
- 5. Model routing selects replaceable local or hosted resources by workload, privacy, latency, cost, and evidence confidence.
- 6. Controlled execution services perform approved actions inside the required isolation tier.
- 7. Independent verification evaluates outputs and observed state.
- 8. Evidence services record the complete decision, execution, verification, and approval chain.
- 9. Human operators approve, interrupt, correct, roll back, or retire consequential work.

Critical trust boundaries

BOUNDARY	PRIMARY THREAT	REQUIRED CONTROL
Human to interface	Spoofed identity, ambiguous intent, unsafe approval, or inaccessible interaction.	Strong authentication, typed confirmation, consequence display, accessibility, and cancellation.
Content to context	Prompt injection, poisoned source, stale authority, or data exfiltration.	Isolation, authority ranking, content labeling, egress policy, and provenance.
Model to tool	Privilege escalation, confused deputy, secret exposure, or malformed action.	Typed contract, capability grant, brokered secret, validation, sandbox, and policy obligations.
Agent to agent	Unbounded delegation, context contamination, identity confusion, or collusion.	Typed delegation packet, principal identity, scope, budget, provenance, and supervisor policy.
Runtime to host	Process escape, resource exhaustion, persistence, or unauthorized network access.	OS and workload isolation, quotas, signed artifacts, network controls, and cleanup.
State to evidence	Tampering, missing events, replay ambiguity, or inconsistent final state.	Durable event identity, hashes, signatures, ordering, observed-state verification, and retention.
Local to hosted	Sovereignty breach, provider retention, traffic analysis, or service dependency.	Policy-controlled routing, minimization, encryption, contractual controls, and tested fallback.
Extension ecosystem	Malicious package, update compromise, excessive permissions, or dependency confusion.	Publisher identity, signing, provenance, review, trust tier, permissions, and revocation.

Deployment profiles

- Sovereign local-first workstation or cluster for protected inference and durable state.
- Enterprise hybrid model fabric with centralized identity, policy, observability, egress, and evidence.
- Disconnected high-assurance environment with signed offline updates and restricted extensions.
- Cloud-managed platform with externally controlled identity, policy, evidence, data governance, and exit strategy.
- Developer and evaluation environment with synthetic data, bounded credentials, reproducible tools, and no production authority.

Failure and recovery behavior

FAILURE EVENT	DESIRED SYSTEM BEHAVIOR	VERIFICATION
Planner or supervisor crash	Resume from durable checkpoint without silently repeating irreversible actions.	Kill test during every mission phase.
Model provider unavailable	Route to validated fallback or suspend safely while preserving state and evidence.	Provider outage and latency fault injection.
Policy service unavailable	Fail closed for consequential actions; allow only explicitly safe degraded operations.	Dependency isolation and deny-path test.
Credential or identity compromise	Revoke principal and grants, contain sessions, rotate secrets, and trace affected actions.	Revocation propagation and forensic reconstruction.
Tool produces unexpected side effect	Interrupt, contain, reconcile observed state, and invoke rollback or manual recovery.	Faulted tool and partial-commit scenarios.
State store divergence	Preserve competing versions, block unsafe promotion, and execute explicit reconciliation.	Offline edits and partition recovery.
Verifier disagreement	Escalate uncertainty rather than selecting a convenient result.	Independent evaluator and human arbitration test.
Evidence store degradation	Stop assured promotion when records cannot be retained or verified.	Write failure, tampering, and restore test.

Architecture validation plan

VALIDATION DOMAIN	ACCEPTANCE TEST
Identity and authority	Trace every action to a principal, grant, policy decision, and expiration; prove revocation.
Isolation	Demonstrate denial of undeclared filesystem, network, secret, process, and cross-tenant access.
Durability	Recover missions after process, host, queue, database, model, and tool failure.
Safety and policy	Exercise prohibited actions, approval obligations, rate and budget limits, and emergency stop.
Knowledge integrity	Test poisoned, stale, conflicting, inaccessible, and deleted sources and memories.
Verification	Compare independent methods, observed state, negative tests, and human review.
Evidence	Reconstruct a mission end to end and verify record integrity, time, identity, and final disposition.
Portability	Replace model, provider, vector store, workflow runtime, and extension without losing authoritative state.

Architecture decisions requiring explicit records

- Authority source for mission, identity, policy, state, and evidence.
- Model and provider replacement boundary.
- Context and memory scopes, retention, correction, and deletion.
- Tool trust tiers, permissions, isolation, and secret brokerage.

- Durability, consistency, replay, and irreversible side-effect strategy.
- Verification independence and disagreement handling.
- Evidence format, integrity, retention, access, and disclosure.
- Sovereignty, egress, residency, and disconnected-operation profile.