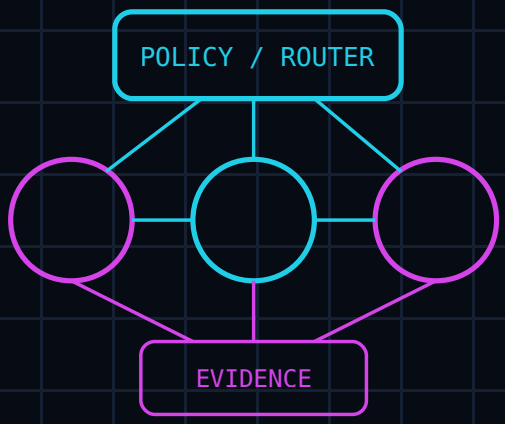



CANDIDATE STANDARD

# Mythos Agentic Systems and Model Orchestration Standard

Bound delegated action with mission contracts, policy, tool controls, memory governance, budgets, evidence, verification, and containment.

PERMANENT ID
MYTHOS-FND-0010

PUBLICATION CLASS
CANDIDATE STANDARD

VERSION / STATUS
1.0.0-candidate / CANDIDATE

PERMANENT PUBLICATION IDENTITY

# Mythos Agentic Systems and Model Orchestration Standard

Universal Requirements for Governed Agents, Multi-Model Routing, Tool Use, Memory, Durable Workflows, Delegation, and Autonomous Execution

Document ID
MYTHOS-FND-0010

Version
1.0.0-candidate

Status
Candidate Standard

Review date
2027-09-17

**Publication position**
This is a permanent candidate standard in the Mythos Engineering Standards Library. Candidate status means the content is ready for structured implementation and review, but it is not represented as external certification, regulatory approval, or independent assurance.

| PROFILE | CONTROLS | CADENCE | EVIDENCE |
|---|---|---|---|
| 3 LEVELS | 23 | TRIGGERED | ASSESSABLE |
| Foundational / Advanced / High Assurance | Atomic normative controls | Annual plus material-change review | Examine / Interview / Test / Measure |

## Reader orientation

| Dimension | Engineering position |
|---|---|
| Primary domain | Foundation and Governance |
| Audience | Agent and AI platform architects; Software, automation, and infrastructure engineers; Security, safety, privacy, and governance teams; Operations and incident-response teams; Assessors and model-risk owners |
| Publisher / owner | Mythos Systems / Standards Program |
| Public classification | Public technical standard |
| Canonical route | /library/standards/foundation/mythos-fnd-0010/ |

## Expected outcomes

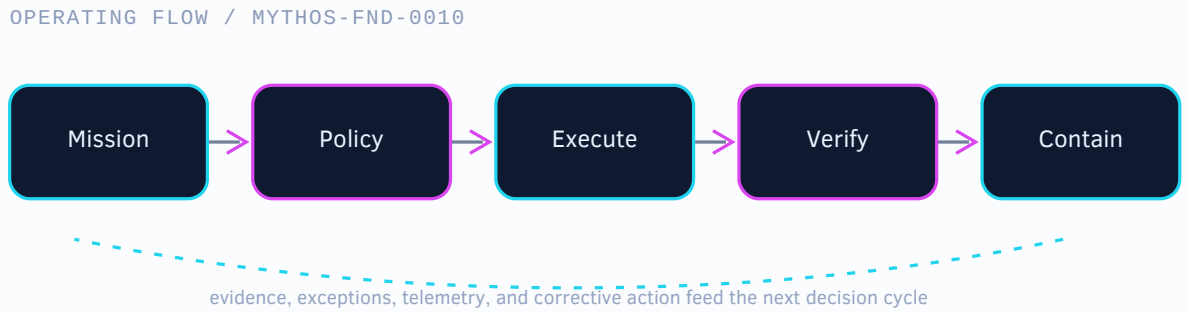
- Separate model reasoning from execution authority and policy enforcement.
- Constrain delegation, recursion, memory, tools, budgets, and cross-agent communication.
- Make every consequential action attributable, reviewable, resumable, and recoverable.
- Enable multi-model and multi-agent systems without surrendering security, reliability, or human accountability.

SOURCE / FRESHNESS POSITION

# Authority and source posture

**Validated 2026-09-17**
Primary authority versions were revalidated for this regeneration on 2026-09-17. Current-source updates are reflected where a newer stable version was identified; active drafts are labeled as such and do not silently replace current final authorities.

FOUNDATION OPERATING DOCTRINE




The source lineage for this edition is the enterprise candidate source set produced in July 2026. Normative controls are preserved; editorial, visual, metadata, and authority-freshness changes are recorded as revision lineage rather than presented as new control substance.

NAVIGATION

# Contents

The table of contents is page-aware and internally linked. Control-level navigation follows on the next pages.

NORMATIVE SYSTEM

# Control index

Every control is independently addressable and assessable. Page references link directly to the control record.

# Document Control

| Field | Value |
|---|---|
| Document ID | MYTHOS-FND-0010 |
| Standard | Agentic Systems and Model Orchestration |
| Version | 1.0.0-candidate |
| Status | Candidate Standard |
| Classification | Public |
| Publisher | Mythos Systems |
| Program | Mythos Engineering Standards Program |
| Accountable Owner | Mythos Systems Standards Program |
| Technical Review | Required |
| Source and Citation Review | Required |
| Assessment Review | Required |
| Accessibility and Publication Review | Required |
| Supersedes | No published edition; evolves source draft 0.2.0 |
| Next Review | 2027-09-17 or earlier on trigger |

## Revision Record

| Version | Date | Status | Material Change |
|---|---|---|---|
| 1.0.0-candidate | 2026-09-17 | Candidate Standard | Permanent-publication regeneration: source-preserving editorial system, richer information design, current authority revalidation, stable public metadata, and release QA. |
| 0.2.0 | 2026-07-18 | Working Draft | Source control standard |
| 1.0.0-candidate | 2026-07-22 | Candidate Standard | Enterprise report architecture, executive direction, target operating model, profiles, maturity, authority register, and source-preserving reconstruction |

# How to Use This Standard

## Executive Reading Path

Read Part I, the target-state architecture in Part III, the profile table in Part V, and the assurance dashboard in Part VI.

## Engineering Reading Path

Read Parts II through V, then use the normative controls in Part IV as implementation and design requirements.

## Assessment Reading Path

Read the conformance requirements in Part II, every applicable control's evidence and assessment procedure in Part IV, and the assessment, maturity, confidence, and traceability provisions in Parts VI and VII.

## Normative and Informative Content

Uppercase **MUST**, **MUST NOT**, **SHOULD**, **SHOULD NOT**, and **MAY** statements are normative when used under the requirements-language provisions of this document. Executive findings, examples, implementation patterns, reference architectures, mappings, and external-authority descriptions are informative unless explicitly incorporated into an adopting organization's binding policy, contract, or regulatory obligation.

FOUNDATION STANDARD / STRUCTURAL PART

# Part I – Executive Direction

## 1. Executive Overview

Agentic systems turn model output into multi-step work involving memory, tools, credentials, files, browsers, terminals, infrastructure, and other agents. This creates a larger authority and failure surface than ordinary AI applications. This standard defines agentic systems as governed distributed systems: missions are bounded, identities are explicit, plans are typed, permissions are external, execution is deterministic, state is durable, results are independently verified, and the full chain is preserved as evidence.

The institutional objective is to govern bounded missions, agent and workload identity, external policy, delegation, planning, tools, sandboxing, memory, context, routing, budgets, durable execution, multi-agent coordination, approvals, skills, evidence, verification, containment, incidents, and retirement. Adoption should produce a controlled operating state in which leadership can understand the material risks, engineering teams can act on explicit requirements, assessors can test implementation and operating effectiveness, and the organization can support every conformance or trust claim with current evidence.

## 2. Executive Findings and Required Direction

| Finding | Enterprise Exposure | Required Direction |
|---|---|---|
| Unbounded mission authority | Agents receive broad objectives without explicit scope, budgets, targets, or stopping conditions. | Require a typed mission contract and capability requests. |
| Agent identity collapse | Human, agent, model, tool, and workload identities are conflated. | Use distinct identities and preserve delegation lineage. |
| Tool ambiguity | Natural-language tool descriptions conceal side effects and failure semantics. | Use versioned typed contracts, risk classification, validation, and post-checks. |
| Memory and context poisoning | Untrusted content is promoted into durable memory or treated as instructions. | Separate trust domains, source memory, require review, and support correction and deletion. |
| False execution guarantees | Distributed retries create duplicates or indeterminate effects while systems claim exactly-once execution. | Use idempotency, effect ledgers, reconciliation, and explicit indeterminate state. |
| Multi-agent consensus mistaken for truth | Several models can agree on the same unsupported conclusion. | Use independent evidence, conflict control, deterministic tests, and verification. |

## 3. Business and Operational Impact

### 3.1 Strategic Impact

This standard converts a discipline that is often dependent on individual expertise into an organization-wide system of ownership, records, controls, review, and evidence. It reduces decision ambiguity and makes material assumptions visible before they become embedded in products or operations.

### 3.2 Delivery and Cost Impact

Adoption requires investment in accountable roles, authoritative records, validation, tooling, and review. That cost should be measured against avoidable rework, delayed releases, incidents, regulatory exposure, duplicated platforms, failed integrations, and unsupported product claims.

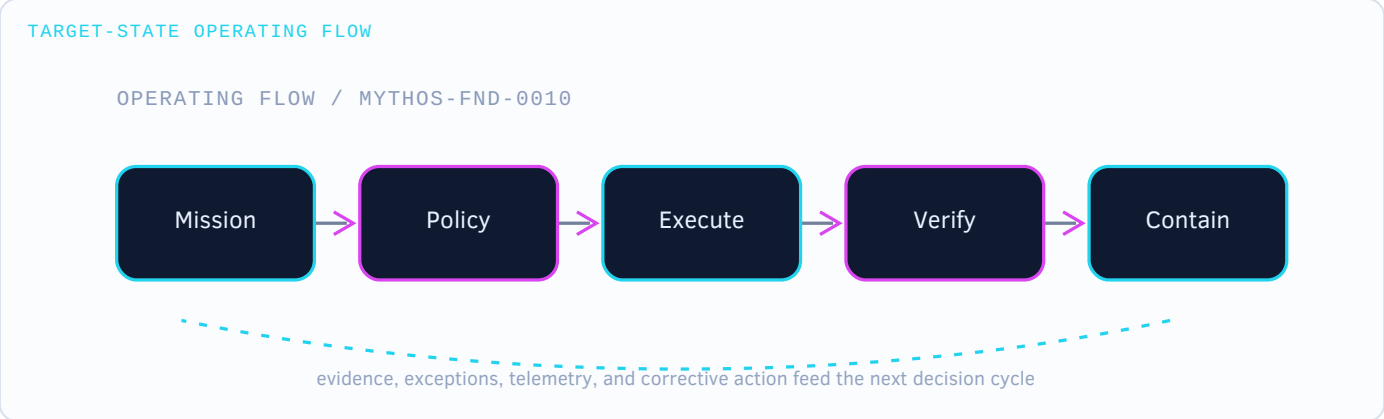
### 3.3 Security, Privacy, and Trust Impact

The standard requires security, privacy, integrity, resilience, and delegated authority to be considered within the primary operating model rather than as downstream approvals. This reduces the likelihood that unsafe boundaries become structurally expensive to correct.

### 3.4 Workforce and Organizational Impact

The organization must distinguish accountable ownership, implementation, approval, and independent challenge. Adoption may expose gaps in authority, skills, evidence, or cross-functional participation that require leadership action.

## 4. Target-State Summary




The target state contains the following institutional components:

| # | Target-State Component |
|---|---|
| 1 | Agent, mission, tool, model, skill, and memory inventory |
| 2 | Bounded mission contract |
| 3 | Human, agent, workload, and tool identity |
| 4 | External policy enforcement |
| 5 | Constrained delegation and lineage |
| 6 | Plan validation and LOGOS-style typed work |
| 7 | Tool contracts and safety classes |
| 8 | Sandboxed and scoped execution |
| 9 | Visual and computer-use control |
| 10 | Governed memory lifecycle |

| # | Target-State Component |
|---|---|
| 11 | Context assembly and trust boundaries |
| 12 | Model routing policy |
| 13 | Resource, time, and recursion budgets |
| 14 | Durable workflow state and effect strategy |
| 15 | Fan-out, fan-in, and conflict resolution |
| 16 | Approvals and separation of duties |
| 17 | Governed skill lifecycle |
| 18 | End-to-end evidence |
| 19 | Independent result verification |
| 20 | Pause, cancel, rollback, and containment |
| 21 | Behavioral and operational monitoring |
| 22 | Agentic incident response |
| 23 | Retirement |

```
Enterprise obligations and risk appetite
    ↓
Accountable ownership and governance
    ↓
Versioned policies, contracts, and authoritative records
    ↓
Engineering implementation and automated enforcement
    ↓
Operational telemetry, evidence, and exception handling
    ↓
Independent assessment and leadership review
    ↓
Corrective action, controlled change, and continuous improvement
```

## 5. Leadership Responsibilities

| Role | Accountability |
|---|---|
| Agentic-system owner | Owns the mission, agent, model, tool, and runtime architecture. |
| Policy authority | Owns capability, delegation, approval, and emergency policy. |
| Runtime owner | Owns durable state, scheduling, retries, cancellation, and effect control. |
| Model-routing owner | Owns model eligibility, fitness, fallback, and cost. |
| Tool and connector owner | Owns contracts, side effects, versioning, and verification. |
| Memory and knowledge owner | Owns context, memory promotion, provenance, correction, and deletion. |
| Independent verifier | Tests plans, outputs, effects, and failure recovery. |
| Incident commander | Coordinates containment, revocation, evidence, and recovery. |

Leadership must ensure that exceptions are explicit, risk-owned, time-bounded, monitored, and retired. A maturity score or successful audit sample does not replace accountability for known nonconformance or residual risk.

FOUNDATION STANDARD / STRUCTURAL PART

# Part II – Standard Foundation

## 6. Purpose and Objectives

- Separate model reasoning from execution authority and policy enforcement.
- Constrain delegation, recursion, memory, tools, budgets, and cross-agent communication.
- Make every consequential action attributable, reviewable, resumable, and recoverable.
- Enable multi-model and multi-agent systems without surrendering security, reliability, or human accountability.

## 7. Scope

Applies to autonomous and semi-autonomous agents, copilots with tools, multi-agent systems, supervisors, planners, model routers, durable agent workflows, computer-use systems, coding agents, infrastructure agents, research agents, and agent-accessible extension ecosystems.

## 8. Intended Audience

- Agent and AI platform architects
- Software, automation, and infrastructure engineers
- Security, safety, privacy, and governance teams
- Operations and incident-response teams
- Assessors and model-risk owners

## 9. Requirements Language and Conformance

The key words **MUST**, **MUST NOT**, **SHOULD**, **SHOULD NOT**, and **MAY** are normative only when written in uppercase and are interpreted in accordance with RFC 2119 and RFC 8174.

A conformance claim MUST identify the assessed scope, standard version, selected assurance profile, tailoring decisions, evidence period, assessor, and unresolved findings. Profiles are cumulative unless a control explicitly states otherwise.

- **Foundational:** broadly applicable minimum controls.
- **Advanced:** stronger governance, automation, scale, and independent verification.
- **High Assurance:** continuous or independent validation, stronger isolation, adversarial testing, formal analysis, or cryptographic evidence where warranted.

## 10. Normative References from Source Draft

- **RFC 2119** — Key words for use in RFCs to Indicate Requirement Levels. https://www.rfc-editor.org/rfc/rfc2119

- **RFC 8174** — Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words. https://www.rfc-editor.org/rfc/rfc8174
- **MYTHOS-GOV-0001** — Mythos Engineering Standards Authoring and Benchmark Framework, Version 0.1.0. ../governance/01_MYTHOS_STANDARDS_AUTHORING_AND_BENCHMARK_FRAMEWORK.md

## 11. Informative References from Source Draft

- **NIST AI 100-1** — Artificial Intelligence Risk Management Framework (AI RMF 1.0). https://doi.org/10.6028/NIST.AI.100-1
- **NIST AI 600-1** — Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. https://doi.org/10.6028/NIST.AI.600-1
- **NIST SP 800-218A** — Secure Software Development Practices for Generative AI and Dual-Use Foundation Models. https://csrc.nist.gov/pubs/sp/800/218/a/final
- **NIST SP 800-53 Rev. 5, Release 5.2.0** — Security and Privacy Controls for Information Systems and Organizations. https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final
- **NIST SP 800-207** — Zero Trust Architecture. https://doi.org/10.6028/NIST.SP.800-207
- **OWASP Top 10 for Agentic Applications 2026** — Risk taxonomy for agentic applications. https://genai.owasp.org/
- **OWASP Top 10 for LLM Applications 2025** — Risk taxonomy for applications using large language models. https://genai.owasp.org/llm-top-10/
- **ISO/IEC 42001:2023** — Artificial intelligence — Management system. https://www.iso.org/standard/81230.html

External mappings are informative unless explicitly incorporated into a contract or regulatory obligation. Adopted organizations remain responsible for validating current source versions and legal applicability.

## 12. Terms and Definitions

| Term | Definition |
|---|---|
| Accountable owner | The person or formally delegated role that accepts responsibility for an outcome, decision, control, or risk. |
| Assessment object | The system, process, component, artifact, team, service, or organizational scope being evaluated. |
| Compensating control | An alternative safeguard that provides comparable risk reduction when the specified control cannot be implemented as written. |
| Conformance claim | A bounded statement that an identified scope satisfies the applicable requirements of a named standard version and assurance profile. |
| Evidence | Reviewable information that supports a conclusion about design, implementation, operation, or control effectiveness. |
| High Assurance | The cumulative profile requiring stronger independence, adversarial testing, continuous verification, or formal evidence where risk warrants it. |
| Residual risk | Risk remaining after controls, mitigations, and compensating measures are applied. |
| System | A product, service, application, platform, workflow, integration, operational capability, or combined socio-technical environment. |
| Tailoring | A documented decision to include, strengthen, parameterize, or mark a control not applicable based on scope and risk. |
| Agent | A software entity that uses one or more models or policies to select, sequence, or adapt actions toward an objective. |

| Term | Definition |
| --- | --- |
| Mission | A bounded unit of agentic work with explicit objective, authority, budget, state, checkpoints, and completion criteria. |
| Tool | A callable capability that can observe, transform, communicate with, or change a resource or environment. |
| Delegation chain | The recorded lineage by which authority, objectives, context, and constraints pass from an originator through supervisors and subagents. |
| Durable state | Persisted workflow state sufficient to resume, inspect, reconcile, or terminate work after interruption. |

## 13. Applicability and Tailoring

The organization MUST determine applicability at the control level. A not-applicable decision MUST identify the assessed scope, factual basis, approver, review date, and any assumptions that would invalidate the decision. Tailoring may strengthen or parameterize a control; it MUST NOT silently weaken a MUST requirement. Compensating controls require documented equivalence of risk reduction.

FOUNDATION STANDARD / STRUCTURAL PART

# Part III – Risk and Architecture

## 14. Enterprise Problem and Risk Landscape

The principal enterprise risks addressed by this standard include inconsistent ownership, undocumented authority, uncontrolled change, local optimization, evidence gaps, stale assumptions, supplier dependence, false assurance, and the inability to determine whether the operating system still matches its approved intent.

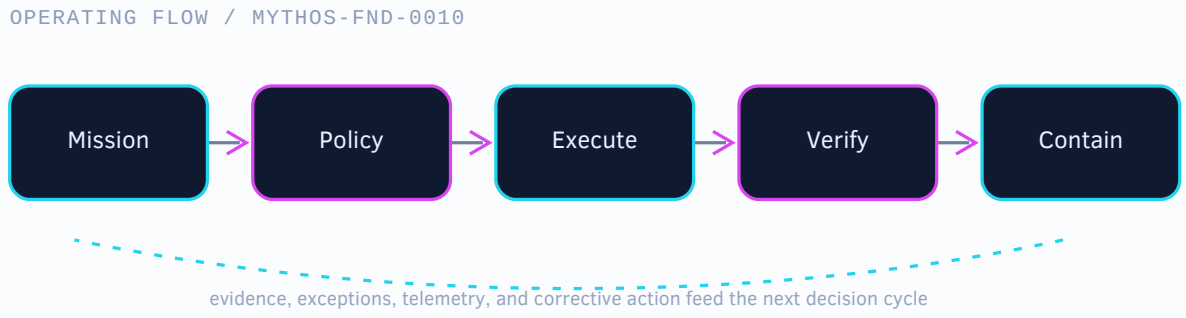
Adopting organizations SHOULD maintain a domain-specific risk register that identifies cause, affected asset or stakeholder, credible scenario, impact, existing controls, residual risk, owner, review trigger, and evidence. Risks that cross standards MUST be linked rather than copied into disconnected registers.

## 15. Governing Principles

- Probabilistic planning must execute through deterministic policy enforcement.
- Delegated authority can narrow but cannot silently expand.
- Every mission has bounded scope, budget, lifetime, and termination conditions.
- Memory is governed data, not an unquestioned source of truth.
- Parallelism and recursion require explicit fan-out, depth, cost, and convergence limits.
- Agents cannot approve their own privilege expansion, evidence sufficiency, or risk acceptance.
- Interruptibility, rollback, reconciliation, and evidence are core runtime properties.

## 16. Reference Operating Architecture

REFERENCE OPERATING ARCHITECTURE



```
Enterprise obligations and risk appetite
    ↓
Accountable ownership and governance
    ↓
Versioned policies, contracts, and authoritative records
    ↓
Engineering implementation and automated enforcement
    ↓
Operational telemetry, evidence, and exception handling
    ↓
Independent assessment and leadership review
    ↓
Corrective action, controlled change, and continuous improvement
```

### 16.1 Control Plane

The control plane contains accountable ownership, policy, standards, risk decisions, approvals, exception governance, and authoritative state. It MUST be distinguishable from implementation and reporting systems.

### 16.2 Delivery and Enforcement Plane

The delivery plane contains engineering workflows, automation, validation, configuration, runtime enforcement, and lifecycle processes. Automated enforcement SHOULD be preferred where requirements can be expressed and tested safely.

### 16.3 Evidence and Assurance Plane

The assurance plane collects evidence, observations, test results, decisions, exceptions, and historical state. Evidence systems MUST protect integrity, scope, provenance, retention, and authorized access.

### 16.4 Feedback Plane

Metrics, incidents, assessments, user outcomes, exceptions, and changing authorities feed corrective action and controlled revision. Feedback MUST not silently alter normative requirements or accepted risk.

## 17. Roles and Accountability

| Role | Accountability |
|---|---|
| Agentic-system owner | Owns the mission, agent, model, tool, and runtime architecture. |
| Policy authority | Owns capability, delegation, approval, and emergency policy. |
| Runtime owner | Owns durable state, scheduling, retries, cancellation, and effect control. |
| Model-routing owner | Owns model eligibility, fitness, fallback, and cost. |
| Tool and connector owner | Owns contracts, side effects, versioning, and verification. |
| Memory and knowledge owner | Owns context, memory promotion, provenance, correction, and deletion. |
| Independent verifier | Tests plans, outputs, effects, and failure recovery. |
| Incident commander | Coordinates containment, revocation, evidence, and recovery. |

## 18. State, Evidence, and Decision Model

The adopting organization MUST distinguish:

- approved intent;
- current implemented state;
- observed operational state;
- generated or inferred recommendations;
- accepted exceptions;
- historical state;
- independently verified results.

A generated report, model conclusion, roadmap, or design proposal MUST NOT be represented as implemented or verified state without supporting evidence.

## 19. Security and Trust Considerations

Every implementation of this standard MUST apply identity, authorization, classification, least privilege, evidence integrity, sensitive-data handling, and supplier-risk controls appropriate to impact. Machine-generated guidance, automation, extensions, and delegated agents MUST remain subject to external policy and independent verification.

## 20. Failure Modes

| Failure or Anti-Pattern | Enterprise Consequence |
|---|---|
| Giving agents a general shell or browser because the user asked for autonomy | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Treating a skill file as permission | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Allowing agents to edit their own policy or evidence | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Using one giant conversation as durable state | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Calling retries exactly-once execution | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Accepting multi-agent agreement without tests | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Allowing computer-use agents unrestricted targets | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |
| Persisting every conversation statement as memory | Failure can create hidden risk, misleading assurance, uncontrolled cost, or unsafe operation. |

## 21. Cross-Functional Dependencies

This Foundation standard depends on the remaining MYTHOS-FND standards for documentation, security and trust, quality, operations, data and integration, interfaces, extensions, AI-first behavior, and agentic orchestration.

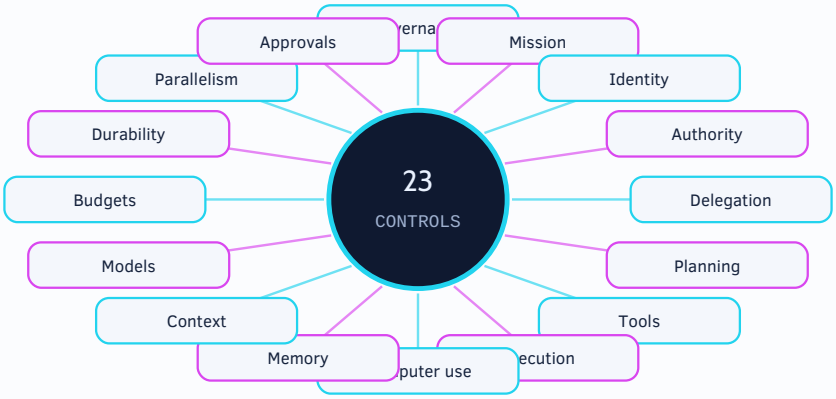
An adopting organization MUST resolve overlapping requirements through the stricter applicable control or a documented authority decision.

FOUNDATION STANDARD / STRUCTURAL PART

# Part IV – Normative Control System

## 22. Control-Family Overview

CONTROL-FAMILY CONSTELLATION




This standard contains **23 controls** across the following families:

- Governance
- Mission
- Identity
- Authority
- Delegation
- Planning
- Tools
- Execution
- Computer use
- Memory
- Context
- Models
- Budgets
- Durability
- Parallelism
- Approvals
- Skills
- Evidence

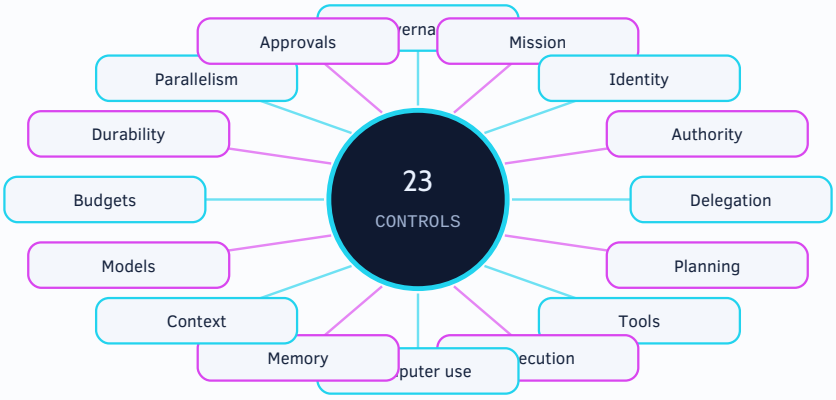
- Verification
- Interruptibility
- Monitoring
- Incident
- Lifecycle

## 23. Normative Controls

| CONTROL SET | PROFILES | ASSESSMENT |
|---|---|---|
| 23 | 3 | 4 |
| atomic requirements | cumulative assurance levels | Examine / Interview / Test / Measure |

Each control is independently addressable, evidence-bound, and assessable. The following control records preserve the source requirements while presenting implementation guidance, required evidence, assessment procedures, exceptions, compensating controls, and external mappings as a consistent engineering system.

NORMATIVE CONTROL CONSTELLATION



NORMATIVE CONTROL

# MYTHOS-FND-AGT-01 – Agent and mission inventory

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Governance | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

The organization MUST inventory agent types, owners, models, tools, memories, environments, identities, permitted missions, autonomy levels, data classes, and maximum consequence.

## Purpose

Creates a bounded governance surface.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Include ephemeral subagents and third-party agents.

## Required evidence

- Agent registry
- Mission catalog
- Ownership records

## Assessment procedure

- **Examine:** Compare runtime identities and tool calls to registry.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- ISO/IEC 42001:2023

NORMATIVE CONTROL

# MYTHOS-FND-AGT-02 – Bounded mission contract

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Mission | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

Every mission MUST define objective, originator, accountable owner, scope, targets, constraints, inputs, authority, budgets, completion criteria, stop conditions, evidence requirements, and expiration before consequential execution.

## Purpose

Prevents open-ended or ambiguous autonomy.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Use machine-readable mission envelopes.
- Reject missing mandatory constraints.

## Required evidence

- Mission record
- Schema validation
- Approval evidence

## Assessment procedure

- **Examine:** Submit ambiguous and expired missions.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST AI RMF 1.0

NORMATIVE CONTROL

# MYTHOS-FND-AGT-03 – Workload and agent identity

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Identity | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

Agents, supervisors, runtimes, tools, and delegated workers MUST use distinguishable, authenticated identities bound to mission and trust-domain context; shared ambient identities MUST NOT authorize consequential actions.

## Purpose

Enables attribution and least privilege.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Use short-lived workload credentials.
- Bind model sessions to execution identity without equating them.

## Required evidence

- Identity architecture
- Credential records
- Authentication tests

## Assessment procedure

- **Examine:** Replay credentials across missions or tenants.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST SP 800-207

NORMATIVE CONTROL

# MYTHOS-FND-AGT-04 – External policy enforcement

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Authority | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

All consequential tool calls and state changes MUST pass through deterministic authorization and policy enforcement outside the model using current identity, mission, resource, tenant, environment, risk, and approval state.

## Purpose

Prevents prompts or model output from self-authorizing.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Default deny.
- Validate arguments and preconditions.

## Required evidence

- Policy rules
- Gateway logs
- Negative tests

## Assessment procedure

- **Examine:** Attempt unauthorized and policy-conflicting calls.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST SP 800-53 — AC family

NORMATIVE CONTROL

# MYTHOS-FND-AGT-05 – Constrained delegation and lineage

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Delegation | Advanced, High Assurance | Critical | High |

## Requirement

Delegation MUST preserve origin, objective, context, constraints, authority, budgets, data classification, and evidence lineage; a delegate MUST NOT expand authority, lifetime, scope, or approval rights beyond the delegator's grant.

## Purpose

Contains authority through agent hierarchies.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Use capability attenuation.
- Record fan-out and delegated sub-budgets.

## Required evidence

- Delegation graph
- Grant records
- Tests

## Assessment procedure

- **Examine:** Attempt privilege amplification and hidden subagents.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST SP 800-53 — AC-6

NORMATIVE CONTROL

# MYTHOS-FND-AGT-06 – Plan validation before execution

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Planning | Foundational, Advanced, High Assurance | Critical | Partial |

## Requirement

Plans for consequential work MUST be validated for scope, dependencies, ordering, preconditions, conflicts, policy, expected effects, rollback, and evidence before execution; validation MUST be independent of the proposing model where risk warrants.

## Purpose

Catches unsafe reasoning before action.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Compile plans to typed operations.
- Use simulation or dry run.

## Required evidence

- Plan schema
- Validation output
- Approval records

## Assessment procedure

- **Examine:** Insert unsafe or contradictory step.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST AI RMF 1.0

NORMATIVE CONTROL

# MYTHOS-FND-AGT-07 – Tool contract and safety classification

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Tools | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

Each tool MUST have a versioned contract defining identity, inputs, outputs, side effects, authorization, idempotency, limits, secrets, failure semantics, observability, reversibility, and risk classification.

## Purpose

Makes tool behavior enforceable.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Separate read from mutate.
- Reject unstructured shell access for routine operations where narrower tools exist.

## Required evidence

- Tool registry
- Schemas
- Conformance tests

## Assessment procedure

- **Examine:** Call with malformed, excessive, and unauthorized arguments.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST SP 800-218A

NORMATIVE CONTROL

# MYTHOS-FND-AGT-08 – Sandboxed and scoped execution

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Execution | Advanced, High Assurance | Critical | High |

## Requirement

Agent-generated code, commands, browser actions, files, and workloads MUST execute in an isolated, resource-bounded, policy-controlled environment appropriate to consequence, with explicit promotion gates to shared or production systems.

## Purpose

Contains malicious or defective execution.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Use disposable environments where practical.
- Control network and filesystem mounts.

## Required evidence

- Sandbox design
- Policy
- Escape tests

## Assessment procedure

- **Examine:** Attempt host, network, secret, and tenant boundary escape.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST SP 800-53 — SC-39

NORMATIVE CONTROL

# MYTHOS-FND-AGT-09 – Visual and computer-use action control

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Computer use | Advanced, High Assurance | Critical | Partial |

## Requirement

Agents operating graphical interfaces or remote computers MUST anchor actions to verified target state, detect unexpected UI changes, avoid hidden or ambiguous elements, limit sensitive capture, and require stronger confirmation for consequential interactions.

### Purpose

Addresses brittle and spoofable visual automation.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Prefer semantic APIs over pixels when available.
- Revalidate after navigation or modal changes.

### Required evidence

- Computer-use policy
- Session recordings or event logs
- Adversarial tests

### Assessment procedure

- **Examine:** Introduce overlays, shifted controls, and deceptive prompts.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST AI 600-1

NORMATIVE CONTROL

# MYTHOS-FND-AGT-10 – Governed memory lifecycle

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Memory | Foundational, Advanced, High Assurance | Critical | Partial |

## Requirement

Agent memory MUST have defined purpose, source, provenance, confidence, scope, tenancy, sensitivity, retention, update, conflict, correction, retrieval, and deletion rules and MUST NOT be treated as authoritative without validation.

## Purpose

Prevents stale, poisoned, or cross-context memory from controlling behavior.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Separate episodic, semantic, procedural, and user-controlled memory.
- Protect shared memory writes.

## Required evidence

- Memory schema
- Access policy
- Poisoning and deletion tests

## Assessment procedure

- **Examine:** Insert conflicting and malicious memory.
- **Interview:** Test tenant isolation.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST AI 600-1

NORMATIVE CONTROL

# MYTHOS-FND-AGT-11 – Context assembly and trust boundaries

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Context | Foundational, Advanced, High Assurance | Critical | Partial |

## Requirement

The runtime MUST label and preserve trust, source, freshness, sensitivity, and instruction status across prompts, retrieved content, memory, tool output, user input, and system policy, and MUST prevent untrusted content from silently becoming higher-priority instruction.

### Purpose

Mitigates context confusion and indirect prompt injection.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Use typed context segments.
- Apply least-context and access filtering.

### Required evidence

- Context schema
- Assembly traces
- Injection tests

### Assessment procedure

- **Examine:** Embed instructions in documents and tool output.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- OWASP Top 10 for Agentic Applications 2026

NORMATIVE CONTROL

# MYTHOS-FND-AGT-12 – Model routing policy

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Models | Advanced, High Assurance | Critical | High |

## Requirement

Model selection and routing MUST follow versioned policy based on task, data sensitivity, capability, risk, latency, cost, availability, jurisdiction, evaluation evidence, and provider terms, with logged rationale and bounded fallback.

## Purpose

Prevents arbitrary or unsafe provider switching.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Never route sensitive context to an unapproved provider.
- Evaluate fallbacks independently.

## Required evidence

- Routing policy
- Model registry
- Routing logs

## Assessment procedure

- **Examine:** Force provider outage and sensitive task.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST AI RMF 1.0

NORMATIVE CONTROL

# MYTHOS-FND-AGT-13 – Resource and recursion budgets

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Budgets | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

Missions MUST enforce approved limits for time, tokens, cost, tool calls, concurrency, subagents, recursion depth, storage, network, and changed resources, with deterministic stop or escalation behavior.

## Purpose

Prevents runaway consumption and uncontrolled scope growth.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Allocate sub-budgets rather than duplicating parent budget.

## Required evidence

- Budget policy
- Runtime counters
- Limit tests

## Assessment procedure

- **Examine:** Trigger each budget limit and race conditions.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- NIST AI 600-1

NORMATIVE CONTROL

# MYTHOS-FND-AGT-14 – Durable state and exactly-once effect strategy

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Durability | Advanced, High Assurance | Critical | High |

## Requirement

Consequential workflows MUST persist authoritative state, action identifiers, approvals, checkpoints, tool outcomes, retries, and compensation status sufficient to resume or reconcile without unintended repeated effects.

### Purpose

Controls interruption and uncertain outcomes.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Separate workflow state from model transcript.
- Use idempotency and deduplication.

### Required evidence

- State model
- Execution journal
- Recovery tests

### Assessment procedure

- **Examine:** Crash between request and acknowledgment.
- **Interview:** Resume and reconcile.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- ISO/IEC 25010:2023 — reliability

NORMATIVE CONTROL

# MYTHOS-FND-AGT-15 – Fan-out, fan-in, and conflict control

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Parallelism | Advanced, High Assurance | Critical | Partial |

## Requirement

Parallel agents and tasks MUST have explicit partitioning, authority, shared-state rules, synchronization, conflict detection, result validation, convergence criteria, and bounded fan-out.

## Purpose

Prevents duplicated effects and inconsistent synthesis.

## Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

## Implementation guidance

- Use immutable inputs where possible.
- Do not allow competing writers without coordination.

## Required evidence

- Concurrency design
- Task graph
- Conflict tests

## Assessment procedure

- **Examine:** Run duplicate and conflicting agents.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

## Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

## External mappings

- ISO/IEC 25010:2023

NORMATIVE CONTROL

# MYTHOS-FND-AGT-16 – Independent approval and separation of duties

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Approvals | Foundational, Advanced, High Assurance | Critical | High |

## Requirement

Agents MUST NOT approve their own privilege expansion, production promotion, evidence sufficiency, policy exception, risk acceptance, or irreversible high-consequence action; required approval MUST come from an authorized independent principal or deterministic policy.

### Purpose

Preserves accountability and prevents self-authorization.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Define quorum or dual control for highest risk.

### Required evidence

- Approval policy
- Decision records
- Bypass tests

### Assessment procedure

- **Examine:** Attempt self-approval and approval replay.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST SP 800-53 — AC-5

NORMATIVE CONTROL

# MYTHOS-FND-AGT-17 – Governed skill creation and modification

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Skills | Advanced, High Assurance | Critical | Partial |

## Requirement

Agent-created or modified skills, prompts, policies, code, and tools MUST be sandboxed, versioned, evaluated, provenance-tracked, permission-scoped, reviewed, and rollback-capable before broader use; creation MUST NOT imply publication or privilege.

### Purpose

Allows learning without silent self-modification.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Separate propose, test, approve, publish, and execute.

### Required evidence

- Skill registry
- Provenance
- Evaluation and review records

### Assessment procedure

- **Examine:** Create a skill requesting new privilege.
- **Interview:** Attempt silent replacement.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST SP 800-218A
- SLSA 1.2

NORMATIVE CONTROL

# MYTHOS-FND-AGT-18 – End-to-end execution evidence

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Evidence | Advanced, High Assurance | Critical | High |

## Requirement

Consequential missions MUST produce tamper-evident evidence linking objective, inputs, context sources, plan, models, prompts or policy versions, delegation, approvals, tool calls, diffs, outputs, verification, exceptions, and final state.

### Purpose

Supports audit and trustworthy completion claims.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Store hashes or references instead of unnecessary sensitive payloads.

### Required evidence

- Evidence bundle
- Integrity validation
- Reconstruction report

### Assessment procedure

- **Examine:** Reconstruct a sampled mission and detect tampering.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST SP 800-53 — AU family

NORMATIVE CONTROL

# MYTHOS-FND-AGT-19 – Independent result verification

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Verification | Foundational, Advanced, High Assurance | Critical | Partial |

## Requirement

Agents MUST verify consequential outcomes against explicit acceptance criteria using authoritative observations, tests, reconciliation, or independent evaluators before declaring completion.

### Purpose

Prevents self-reported success from substituting for reality.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Use different evidence paths than the action path where possible.

### Required evidence

- Acceptance criteria
- Verification results
- Completion decision

### Assessment procedure

- **Examine:** Cause partial failure and misleading tool output.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST AI RMF 1.0

NORMATIVE CONTROL

# MYTHOS-FND-AGT-20 – Pause, cancel, rollback, and containment

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Interruptibility | Advanced, High Assurance | Critical | High |

## Requirement

The runtime MUST support authorized pause, cancellation, kill, isolation, rollback or compensation, and evidence preservation for active missions, including cascaded subagents and pending tool work.

### Purpose

Enables human and automated containment.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Kill controls must not depend solely on the affected agent.

### Required evidence

- Control design
- Drill results
- State reconciliation

### Assessment procedure

- **Examine:** Cancel during fan-out and partial execution.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST SP 800-61 Rev. 3

NORMATIVE CONTROL

# MYTHOS-FND-AGT-21 – Behavioral and operational monitoring

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Monitoring | Advanced, High Assurance | Critical | Partial |

## Requirement

Agent runtimes MUST monitor policy denials, novel tool sequences, authority use, budget consumption, recursion, delegation, errors, drift, unsafe content, anomalous data access, and outcome quality with actionable escalation.

### Purpose

Detects emergent and compromised behavior.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Baseline by agent and mission type.
- Avoid relying on model self-report.

### Required evidence

- Telemetry plan
- Detections
- Incident records

### Assessment procedure

- **Examine:** Generate anomalous sequences and confirm alerting.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST SP 800-53 — SI-4

NORMATIVE CONTROL

# MYTHOS-FND-AGT-22 – Agentic incident response

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Incident | Advanced, High Assurance | Critical | Partial |

## Requirement

The organization MUST maintain incident procedures for compromised agents, poisoned memory or context, malicious tools, provider failure, runaway missions, unauthorized actions, evidence compromise, and cross-tenant exposure.

### Purpose

Addresses the distinctive failure modes of agentic systems.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Support global and scoped revocation.
- Preserve mission state.

### Required evidence

- Incident plan
- Exercises
- Corrective actions

### Assessment procedure

- **Examine:** Run a memory-poisoning or runaway-agent exercise.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- NIST SP 800-61 Rev. 3

NORMATIVE CONTROL

# MYTHOS-FND-AGT-23 – Agent, model, tool, and memory retirement

| FAMILY | PROFILES | FAILURE IMPACT | AUTOMATION POTENTIAL |
|---|---|---|---|
| Lifecycle | Advanced, High Assurance | Critical | High |

## Requirement

Retirement MUST revoke identities and capabilities, terminate schedules and missions, migrate or delete governed state and memory, preserve required evidence, update dependents, and verify no residual execution path remains.

### Purpose

Prevents orphaned autonomous authority.

### Applicability

Applicable to all in-scope systems and organizational processes unless a documented tailoring decision states otherwise.

### Implementation guidance

- Search for hidden schedules, webhooks, and credentials.

### Required evidence

- Retirement plan
- Revocation evidence
- Dependency report

### Assessment procedure

- **Examine:** Retire a test agent and attempt delayed execution.
- **Interview:** Confirm ownership and practice.
- **Test:** Exercise the requirement.

### Exceptions and compensating controls

Any exception must be approved by an accountable owner, time-bounded, risk-assessed, monitored, and supported by compensating controls.

### External mappings

- ISO/IEC/IEEE 15288:2023

FOUNDATION STANDARD / STRUCTURAL PART

# Part V – Implementation and Operations

## 24. Implementation Profiles

| Profile | Required Operating State |
|---|---|
| Baseline | Inventory, bounded mission, distinct identity, external policy, validated plan, typed tools, scoped execution, governed memory and context, routing, budgets, durable state, approval, evidence, verification, cancellation, monitoring, incident response, and retirement. |
| Enterprise | Central agent registry, multi-agent runtime, policy-as-code, worker attestation, context compiler, model routing service, skill governance, continuous observability, evidence graph, and enterprise incident operations. |
| High Assurance | Formally constrained missions, capability-secure execution, microVM or equivalent isolation, independent plan review, two-person approval for high-risk actions, deterministic effect ledger, continuous verification, adversarial agent testing, and cryptographically protected evidence. |

Profiles are cumulative unless a control explicitly states otherwise. A higher profile cannot be claimed solely through additional tooling; governance, evidence, operating effectiveness, and independent challenge must also satisfy the profile.

## 25. Implementation Lifecycle

### 25.1 Establish

- appoint accountable ownership;
- determine applicability and profile;
- inventory current processes, systems, records, and known exceptions;
- identify authority sources and legal applicability;
- establish evidence locations and review cadence.

### 25.2 Baseline

- assess every applicable control;
- distinguish implemented, partially implemented, not implemented, not applicable, and not assessed;
- record evidence quality and confidence;
- identify systemic dependencies and priority risks.

### 25.3 Implement

- define target architecture and ownership;
- create or revise policies, contracts, workflows, and controls;
- automate enforcement and evidence where safe;
- test failure and recovery paths;
- train accountable roles.

### 25.4 Assure

- conduct technical and procedural assessment;
- verify evidence over a representative period;
- test sampled controls and critical workflows;
- challenge exceptions, unsupported claims, and optimistic assumptions.

### 25.5 Operate and Improve

- monitor metrics and exceptions;
- investigate incidents and control failures;
- update for authority, threat, technology, or organizational change;
- preserve superseded decisions and evidence;
- retire obsolete controls and implementations deliberately.

## 26. Integration Requirements

Implementations SHOULD integrate the standard with product governance, architecture review, secure development, CI/CD, change management, observability, service management, identity, evidence, risk, procurement, and training systems. Integration MUST preserve ownership and source authority rather than copying uncontrolled state between tools.

## 27. Failure Handling and Recovery

Control failures MUST produce a classified finding, owner, containment or compensating action, due date, evidence requirement, and escalation path. Where failure creates ongoing material risk, the organization MUST consider stopping, restricting, rolling back, isolating, or retiring the affected activity.

## 28. Exceptions and Compensating Controls

Exceptions MUST identify the exact control, scope, rationale, risk, owner, approval, compensating measures, monitoring, expiration, and closure evidence. Exceptions MUST NOT be used to convert a permanent design weakness into nominal conformance.

## 29. Prohibited Practices

- Giving agents a general shell or browser because the user asked for autonomy
- Treating a skill file as permission
- Allowing agents to edit their own policy or evidence
- Using one giant conversation as durable state
- Calling retries exactly-once execution
- Accepting multi-agent agreement without tests
- Allowing computer-use agents unrestricted targets
- Persisting every conversation statement as memory

## 30. Adoption Roadmap from Source Draft

- Define mission, agent, tool, policy, identity, budget, state, and evidence schemas.
- Route all tool use through an authenticated, policy-enforced execution gateway.
- Add durable orchestration, checkpoints, approvals, cancellation, reconciliation, and audit.
- Govern memory, delegation, multi-model routing, skills, and extension access.
- For High Assurance, add independent supervisors, formal policy constraints, isolated execution ranges, adversarial simulation, and continuous verification.

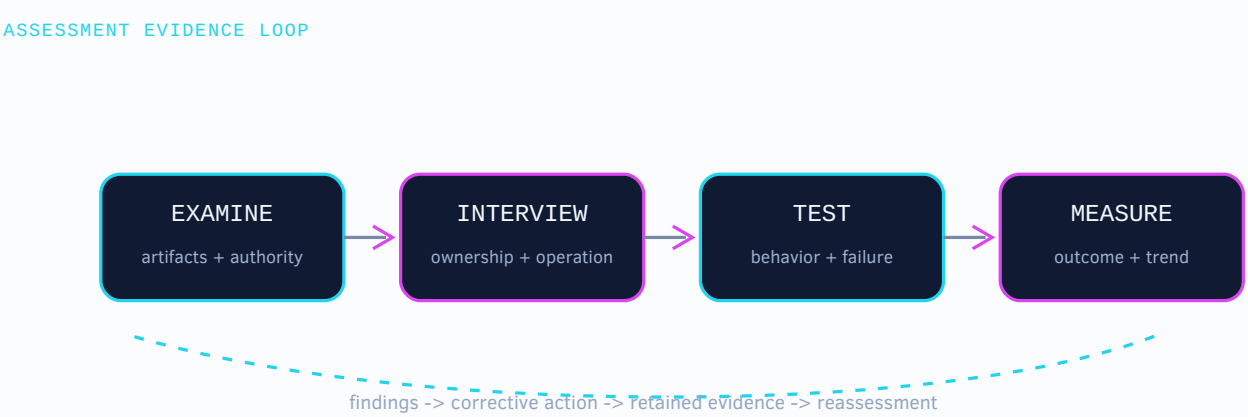
## 31. Limitations and Residual Risk from Source Draft

Conformance demonstrates that defined requirements were satisfied within a bounded scope and evidence period. It does not guarantee absence of defects, compromise, harm, outage, legal exposure, or future failure. Novel threats, misunderstood dependencies, invalid assumptions, human error, supplier changes, and conditions outside the assessment scope remain possible.

FOUNDATION STANDARD / STRUCTURAL PART

# Part VI – Assurance and Assessment

## 32. Assessment Methodology




Assessors MUST use a combination of examination, interview, and testing appropriate to the selected profile and control risk. Assessment records MUST identify sample size, tools, versions, evidence references, exclusions, limitations, confidence, and residual uncertainty. Automated checks MAY support a finding but MUST NOT be treated as proof of effective governance or operation when the control depends on human accountability or contextual judgment.

### 12.1 Finding States

- **Satisfied:** requirement implemented and supported by sufficient evidence.
- **Partially satisfied:** some required outcomes or scope are missing.
- **Not satisfied:** requirement absent, ineffective, or contradicted by evidence.
- **Not applicable:** supported by an approved tailoring rationale.
- **Not assessed:** evidence is insufficient.

### 12.2 Conformance

A scope is **Conformant** only when all applicable MUST controls for the claimed profile are satisfied. A failed critical MUST control cannot be averaged away by stronger performance elsewhere.

## 33. Metrics and Reporting from Source Draft

| Metric | Definition | Expected use or target |
|---|---|---|
| Attributed action coverage | Consequential actions linked to mission, actor, model, policy, tool, and evidence | 100% |
| | Actions outside granted authority reaching execution | |

| Metric | Definition | Expected use or target |
|---|---|---|
| Unauthorized action rate | | 0 in defined tests and production |
| Mission reconciliation | Interrupted missions reconciled to known state | 100% of consequential missions |
| Budget enforcement | Missions exceeding approved cost, time, token, tool, or concurrency budgets without approval | 0 |
| Evaluation pass rate | Release-critical agent scenarios meeting deterministic acceptance rules | 100% of MUST scenarios |

Metrics MUST be interpreted with scope and uncertainty. Organizations SHOULD report trends, distributions, and exceptions rather than presenting a single composite score as complete assurance.

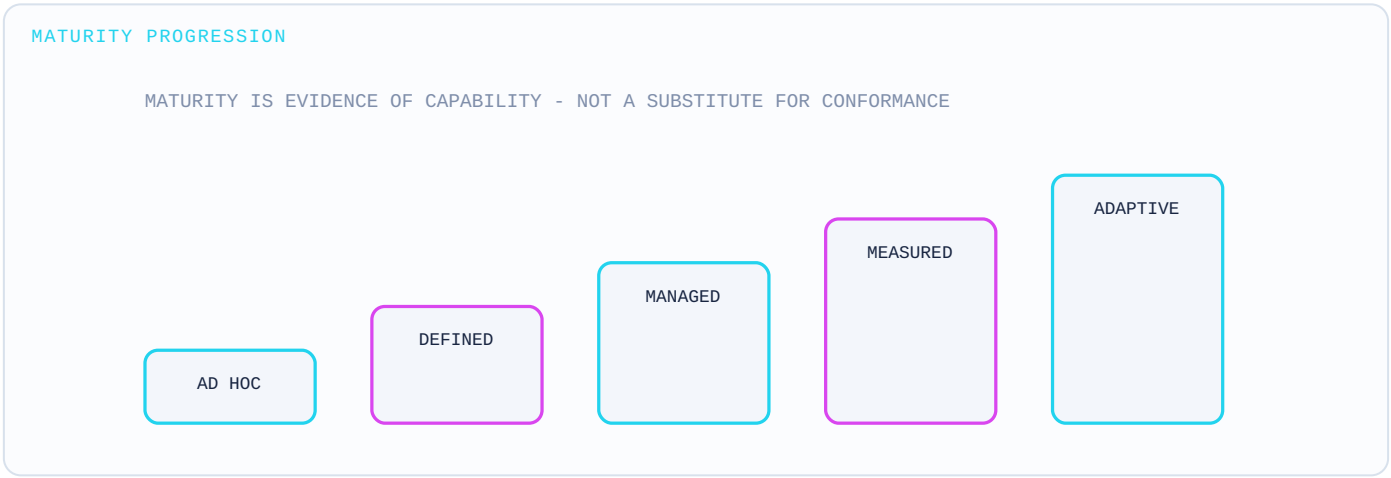
## 34. Exceptions and Compensating Controls

Every exception MUST identify the affected control and scope, justification, risk, compensating controls, accountable approver, start and expiration dates, monitoring, and remediation or retirement plan. Exceptions MUST be reevaluated after material change or incident and MUST NOT be self-approved by the team or agent requesting the exception where separation of duties is required.

## 35. Enterprise Metric Set

| Metric | Measurement Requirement |
|---|---|
| missions without bounded contract | Defined by the adopting organization with owner, source, target, and cadence. |
| agents without distinct workload identity | Defined by the adopting organization with owner, source, target, and cadence. |
| tool calls outside declared contract | Defined by the adopting organization with owner, source, target, and cadence. |
| delegation depth and unresolved lineage | Defined by the adopting organization with owner, source, target, and cadence. |
| context sources without trust classification | Defined by the adopting organization with owner, source, target, and cadence. |
| memory proposals accepted without review | Defined by the adopting organization with owner, source, target, and cadence. |
| model and tool calls beyond budget | Defined by the adopting organization with owner, source, target, and cadence. |
| indeterminate effects awaiting reconciliation | Defined by the adopting organization with owner, source, target, and cadence. |
| high-risk actions without independent approval | Defined by the adopting organization with owner, source, target, and cadence. |
| results lacking independent verification | Defined by the adopting organization with owner, source, target, and cadence. |
| time to revoke and contain compromised agent | Defined by the adopting organization with owner, source, target, and cadence. |

# 36. Maturity Model




| Level | Name | Required Characteristics |
|---|---|---|
| M0 | Unmanaged | Outcome is not intentionally controlled or evidence is unavailable. |
| M1 | Documented | Ownership and procedures exist but implementation is inconsistent or mostly manual. |
| M2 | Implemented | Applicable controls operate in the assessed scope and evidence exists. |
| M3 | Measured | Effectiveness, exceptions, failures, and trends are measured and reviewed. |
| M4 | Assured | Independent testing, representative evidence, and continuous or periodic validation exist. |
| M5 | Adaptive | Controls respond safely to changing risk, authority, technology, and operational evidence. |

Maturity does not override conformance. An organization cannot compensate for a failed mandatory requirement by assigning itself a high maturity level.

# 37. Assessment Confidence

Assessment confidence MUST be recorded as Low, Moderate, High, or Very High. Confidence considers evidence integrity, observation period, sample coverage, source authority, environmental representativeness, test repeatability, reviewer independence, and unresolved gaps.

# 38. Executive Assurance Dashboard

The executive report SHOULD present:

- applicable controls;
- conformant, partial, nonconformant, exception, not applicable, and not assessed counts;
- high and critical findings;
- evidence freshness;
- exception age;
- maturity distribution;
- assessment confidence;
- top residual risks;

- corrective-action status.

FOUNDATION STANDARD / STRUCTURAL PART

# Part VII – Authority and Traceability

## 39. Cross-Standard Dependencies from Source Draft

This standard is part of an integrated foundation. Product decisions depend on documentation and traceability; implementation depends on security, quality, data, interfaces, and extension controls; release depends on operational readiness; AI and agentic behavior inherit every applicable non-AI requirement. A specialized standard MAY strengthen these requirements but MUST NOT weaken them without an explicit supersession rule.

## 40. Current External Authority Register

| Authority | Relevance | Official Location | Status | Validated |
|---|---|---|---|---|
| NIST AI RMF 1.0 | AI risk management | https://www.nist.gov/itl/ai-risk-management-framework | Current; revision program active | 2026-09-17 |
| NIST AI 600-1 | Generative AI Profile | https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence | Final | 2026-09-17 |
| NIST SP 800-218A | Secure development for generative AI and dual-use foundation models | https://csrc.nist.gov/pubs/sp/800/218/a/final | Final | 2026-09-17 |
| NIST SP 800-207 | Zero Trust Architecture | https://csrc.nist.gov/pubs/sp/800/207/final | Final | 2026-09-17 |
| ISO/IEC 42001:2023 | AI management systems | https://www.iso.org/standard/42001 | Published | 2026-09-17 |
| ISO/IEC 23894:2023 | AI risk management | https://www.iso.org/standard/77304.html | Published | 2026-09-17 |
| ISO/IEC 5338:2023 | AI system life cycle processes | https://www.iso.org/standard/81118.html | Published | 2026-09-17 |
| Regulation (EU) 2024/1689 | Artificial Intelligence Act | https://eur-lex.europa.eu/eli/reg/2024/1689/ | In force; staged application | 2026-09-17 |

The authority register is informative unless an adopting organization incorporates a source into a binding obligation. Legal applicability and licensed ISO content must be evaluated by qualified parties. The register records current source identity and relevance without reproducing protected standards text.

## 41. Control Traceability

Each control MUST remain traceable to:

- the risk or outcome it addresses;
- the applicable profile;
- implementation ownership;
- evidence;

- assessment procedure;
- exceptions;
- external mappings;
- related Foundation controls.

## 42. Source-Draft Preservation

This enterprise candidate edition preserves every normative control and the substantive source-draft sections. Editorial movement into the enterprise report structure does not remove the original requirement, implementation guidance, evidence, assessment procedure, exception rule, or external mapping.

FOUNDATION STANDARD / STRUCTURAL PART

# Part VIII – Appendices

## Appendix A – Source-Draft Evolution Notes

- Preserves the source draft's model-independent harness, hierarchy, memory, fan-out/fan-in, skills, budgets, and evidence concepts.
- Reframes product-specific names as universal runtime roles and control points.
- Adds explicit identity, delegation lineage, durable-state reconciliation, computer-use, and autonomous-change controls.

## Appendix B – Change History

| Version | Date | Status | Summary |
|---|---|---|---|
| 0.2.0 | 2026-07-18 | Working Draft | First standards-program restructuring of the source draft into atomic controls, assurance profiles, evidence, assessment procedures, and cross-standard dependencies. |

## Appendix C – Control Summary

| Control | Family | Title | Profiles | Automation | Impact |
|---|---|---|---|---|---|
| MYTHOS-FND-AGT-01 | Governance | Agent and mission inventory | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-02 | Mission | Bounded mission contract | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-03 | Identity | Workload and agent identity | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-04 | Authority | External policy enforcement | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-05 | Delegation | Constrained delegation and lineage | Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-06 | Planning | Plan validation before execution | Foundational, Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-07 | Tools | Tool contract and safety classification | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-08 | Execution | Sandboxed and scoped execution | Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-09 | Computer use | Visual and computer-use action control | Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-10 | Memory | Governed memory lifecycle | Foundational, Advanced, High Assurance | Partial | Critical |
| | Context | | | Partial | Critical |

| Control | Family | Title | Profiles | Automation | Impact |
|---|---|---|---|---|---|
| MYTHOS-FND-AGT-11 | | Context assembly and trust boundaries | Foundational, Advanced, High Assurance | | |
| MYTHOS-FND-AGT-12 | Models | Model routing policy | Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-13 | Budgets | Resource and recursion budgets | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-14 | Durability | Durable state and exactly-once effect strategy | Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-15 | Parallelism | Fan-out, fan-in, and conflict control | Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-16 | Approvals | Independent approval and separation of duties | Foundational, Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-17 | Skills | Governed skill creation and modification | Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-18 | Evidence | End-to-end execution evidence | Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-19 | Verification | Independent result verification | Foundational, Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-20 | Interruptibility | Pause, cancel, rollback, and containment | Advanced, High Assurance | High | Critical |
| MYTHOS-FND-AGT-21 | Monitoring | Behavioral and operational monitoring | Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-22 | Incident | Agentic incident response | Advanced, High Assurance | Partial | Critical |
| MYTHOS-FND-AGT-23 | Lifecycle | Agent, model, tool, and memory retirement | Advanced, High Assurance | High | Critical |

# Appendix D – Minimum Conformance Claim Record

```
standard: MYTHOS-FND-0010
version: 0.2.0
profile: foundational | advanced | high_assurance
scope: <bounded systems, teams, environments, and processes>
assessment_period: <start/end>
assessor: <person or independent body>
tailoring_decisions: []
exceptions: []
findings_summary:
  satisfied: 0
  partially_satisfied: 0
  not_satisfied: 0
  not_applicable: 0
  not_assessed: 0
residual_risk_owner: <accountable role>
approval: <record reference>
```

# Appendix E – Publication Review Checklist

- [ ] Domain technical review completed

- [ ] Security and privacy review completed
- [ ] Current primary sources validated
- [ ] Exact external mappings reviewed
- [ ] All controls testable
- [ ] Evidence requirements sufficient
- [ ] Assessment procedures repeatable
- [ ] Executive findings supported
- [ ] Legal applicability reviewed where required
- [ ] Accessibility and publication quality verified
- [ ] Machine-readable control catalog validated
- [ ] Change and review record complete
- [ ] Formal approval recorded