ODA3-2026-08-WHP-COM-020

ODA3 INSTITUTE

How to Assess an AI System

Examination Perspective

WHP ADOPTION GUIDE

Public Version 1.1

PUBLISHED: AUGUST 2026

ODA3 Institute • www.oda3.org

Publication Notice

This is Public Version 1.1 of ODA3-2026-08-WHP-COM-020, the Final Publication, issued 28 October 2026. It replaces Version 1.0 (August 2026). Public Version 1.1 adds the licence and copyright notice, adds §21.1 (boundary changes), which earlier versions already cited, and removes production notes. The method is otherwise unchanged. It provides a practitioner methodology for organizations examining their own AI systems. It does not describe or activate an ODA3 assessment or certification programme, grant assessment authority, or establish conformity, certification, functional safety, or an independent assurance conclusion.

Publication boundary: Findings remain bounded to the stated system, scope, period, conditions, evidence basis and procedure. Supported / Partial / Gap / Unknown are descriptive finding aids as defined in §81, not scores, severity grades, certification states or universal decision thresholds.

© 2026 ODA3 Pvt Ltd. Published by ODA3 Institute. Classification: Public — GAISSF Ecosystem Licence (GEL) v1.0. GAISSF Ecosystem developed by ODA3 Institute, licensed under GEL v1.0.

Contents

PART I — ASSESSMENT FOUNDATIONS....6
1. What This Guide Is....6
2. What This Guide Is Not....6
3. Why Assessment Needs a Method....6
4. The Central Method....6
5. Recurring Examination Dimensions....7
6. Assurance Language: What a Finding Can and Cannot Say....8
7. Independence and Conflicts....8
8. Relationship to Other ODA3 Practitioner Guides....8
9. What "Notably Absent" Means in This Guide....8
PART II — DEFINE THE ASSESSMENT QUESTION....10
10. Why the Question Comes Before the System....10
11. What Makes a Question Assessable....10
12. The Seven Elements of an Assessment Question Record....10
13. Worked Example: From a Vague Concern to an Assessable Question....11
14. Multiple Questions, One Assessment....11
15. When the Question Changes Mid-Assessment....11
16. Resource-Planning Prompts for This Part....12
PART III — ESTABLISH THE SYSTEM BOUNDARY....13
17. Why the Boundary Comes After the Question....13
18. What Makes an AI-System Boundary Different....13
19. Identify the Components Material to the Question....14
20. Preserve Implementation Responsibility and Evidence Source Separately....14
21. Shared Components and Multi-System Boundaries....15
21.1 When the Boundary Changes During Examination....15
22. Worked Example — Establishing a Boundary....15
23. Physical Consequence Paths....16
24. Practitioner Field — System Boundary Record....17
PART IV — DETERMINE APPLICABILITY....19
25. Why Applicability Must Follow the Boundary....19
26. What "Applicability" Means for an AI Assessment....19
27. Identifying the Applicable Control Baseline....19
28. Mapping Requirements to Candidate System Elements....20
29. Exclusions, Dependencies and Applicability Dispositions....20
29.1 When Applicability Changes During Examination....21
30. Regulatory Context as Applicability Input, Not Determination....21
31. Shared Responsibility and Applicability....22
32. Worked Example — Determining Applicability (Illustrative)....22
33. Practitioner Field — Applicability Record Template (Blank)....23
PART V — UNDERSTAND THE CONTROL....25
34. Why Understanding the Control Comes After Applicability....25
35. What "Understanding the Control" Means....25
36. Identify the Intended Outcome....26
37. Identify the Represented Mechanism....26
38. Ownership and Responsibility....26
39. Intended Operating Boundary....27
40. Hypothesized Failure Conditions....27
40.1 When the Control Understanding Changes During Examination....28
41. Worked Example — Understanding a Control (Illustrative)....28
42. Practitioner Field — Control Understanding Record Template (Blank)....30
PART VI — BUILD THE EVIDENCE BASIS....33

43. Why the Evidence Basis Comes After Control Understanding....33
44. What "Building the Evidence Basis" Means....33
45. Evidence Classification: Design, Implementation, Operating....34
46. Evidence Sources and Provenance....35
47. Temporal Scope, Relationship, and Currency....36
48. Coverage, Missing Evidence, and Inaccessible Evidence....36
48.1 When the Evidence Basis Changes During Examination....37
49. Evidence Preservation and Integrity....37
50. Worked Example — Building the Evidence Basis (Illustrative)....38
51. Practitioner Field — Evidence Basis Record Template (Blank)....40
PART VII — VERIFY CONTROL OPERATION....43
52. Why Verification Comes After the Evidence Basis....43
53. What "Verification" Means in This Guide....43
53.1 Verification result language....44
53.2 Verification procedure is not a universal test protocol....44
54. Verify Propositions About Represented Design....44
54.1 Source conflicts in represented design....45
55. Verify Propositions About Implemented State....45
55.1 Relationship between represented design and implemented state....45
56. Verify Propositions About Operation....46
56.1 Observation and testing....46
57. Corroboration, Contradiction, and Alternative Explanations....46
57.1 Testing alternative explanations....47
58. Refusals, Negative Evidence, and Verification Limitations....47
58.1 When Verification Changes During Examination....48
59. Worked Example — Verifying Control Operation (Illustrative)....48
60. Practitioner Field — Verification Record Template (Blank)....50
PART VIII — ASSESS AI-SPECIFIC BEHAVIOUR....53
61. Why AI-Specific Behaviour Requires Separate Examination....53
62. Probabilistic Behaviour and Output Variability....53
63. Retrieval-Augmented Generation (RAG) System Behaviour....53
64. Agentic Systems and Tool Use....54
65. Model, Provider and Relevant Behavioural Change....54
66. Physical Consequence Paths....55
67. Human Oversight Models....55
68. Worked Example — Assessing AI-Specific Behaviour (Illustrative)....56
69. Practitioner Field — AI-Specific Behaviour Examination Record (Blank)....58
PART IX — HANDLE NEGATIVE AND INCONCLUSIVE RESULTS....60
70. Why Negative and Inconclusive Results Require Separate Treatment....60
71. What "Negative and Inconclusive" Means in This Guide....60
72. Preserving Negative and Inconclusive Results....61
73. Distinguishing Negative Evidence States....61
74. Treatment of Contradictory Evidence....61
75. Treatment of Unresolved and Inconclusive Results....62
76. Worked Example — Handling Negative and Inconclusive Results (Illustrative)....62
77. Practitioner Field — Negative and Inconclusive Results Record (Blank)....63
PART X — FORM FINDINGS AND BOUNDED CONCLUSIONS....65
78. Why Findings Come After Verification and Negative-Result Treatment....65
79. What "Findings" Mean in This Guide....65
80. Constructing a Finding....65
81. Applying the Supported / Partial / Gap / Unknown Convention....66
82. The Boundary Between Finding and Judgment....66
83. Worked Example — Forming Findings (Illustrative)....67
84. Practitioner Field — Findings Record (Blank)....68
85. Finding Revision and Handoff....68

PART XI — REMEDIATION AND REASSESSMENT....70
86. Why Remediation and Reassessment Follow Finding Formation....70
87. Treat Remediation as a Change, Not a Finding Closure....70
88. Preserve Historical Findings....71
89. Corrective Evidence Must Re-enter the Assessment Chain....71
90. Determine the Reassessment Boundary....72
91. Point-in-Time, Event-Triggered and Periodic Reassessment....72
92. Practitioner Field: Remediation and Reassessment Record....73
92.1 Worked Example — Reassessing a Finding After Remediation....74
PART XII — PRACTITIONER FIELD MATERIALS....76
93. Purpose of Part XII....76
94. Record Inventory and Cross-References....76
95. Assessment Question Record....76
96. System Boundary Record....77
97. Applicability Record....77
98. Control Understanding Record....77
99. Evidence Basis Record....77
100. Verification Record and Related Records — Consolidated View....77

CONCLUDING APPARATUS — SOURCE REGISTER / METHODOLOGY NOTES / NOTABLY ABSENT / FINAL PUBLICATION STATEMENT....81

PART I — ASSESSMENT FOUNDATIONS

1. What This Guide Is

An AI system assessment is the disciplined examination of a defined AI system against defined requirements and controls, using available evidence, to reach a bounded conclusion about what that evidence does and does not support.

This guide teaches that discipline. It does not perform an assessment for you, and it does not represent any assessment it helps produce as an ODA3 assessment, certification, or conformity determination.

The guide is written for practitioners who need to assess an AI system their own organization operates, procures, or is otherwise accountable for — security architects, AI governance leads, compliance and risk practitioners, internal assurance and audit teams, system owners preparing for external review, and qualified assessment practitioners engaged by an organization to examine its own systems.

2. What This Guide Is Not

This guide does not describe ODA3 Institute's own assessment or certification programme. As of publication, ODA3's live assessment authority and certification issuance are not activated. Nothing in this guide grants, implies, or represents any assessment authority on ODA3's behalf, and completing the methodology described here does not constitute an ODA3 assessment, certification, or independent assurance conclusion.

This distinction matters enough to restate plainly: an organization that follows this guide has conducted its own assessment, using a defensible method. It has not obtained an ODA3 assessment. The two are different things, and confusing them — even informally, even in an internal summary slide — misrepresents what occurred.

3. Why Assessment Needs a Method

An AI system assessment can fail in two different directions, and both are common.

The first failure is assertion without evidence: a system is declared secure, compliant, or well-governed because someone with relevant authority says so, or because a policy document describes the intended state, without anyone examining whether the intended state actually holds in the operating system. A policy that says "access is restricted to authorized users" is not evidence that access is restricted — it is evidence that the policy exists.

The second failure is precision without foundation: a scoring system, a percentage, or a pass/fail threshold gets attached to an assessment, and the appearance of rigor substitutes for the substance of it. A stated "94% compliance score" sounds more defensible than a plain description of what was and wasn't examined — but if the underlying evidence doesn't actually support the number's precision, the number is doing rhetorical work, not analytical work.

This guide is built to avoid both failures. It insists on evidence rather than assertion, and it insists on bounded, evidence-supported conclusions rather than manufactured precision. Where a genuine standard, regulation, or control baseline establishes a specific numeric requirement, that requirement is used as stated and its source is cited. Where no such external requirement exists, this guide does not invent one.

4. The Central Method

Every assessment described in this guide follows the same chain:

> **assessment question → scope → applicable requirement/control → mechanism → evidence → provenance → verification → alternative explanation → finding → limitation → bounded conclusion**

Each link matters, and skipping one is where assessments typically go wrong:

- **Assessment question** — what, specifically, is being asked? "Is the system secure?" is not an assessment question; it has no defined boundary and no way to be answered with evidence. "Does the deployed retrieval-augmented generation pipeline enforce the same document-level access controls at retrieval time that apply to direct document access?" is an assessment question.
- **Scope** — which system, which version, which period, which environment does the question apply to? A finding about a staging environment is not automatically a finding about production.
- **Applicable requirement/control** — what standard, policy, or control baseline makes this question relevant? Assessing against nothing in particular produces an opinion, not a finding.
- **Mechanism** — what actually implements the control? Not the policy that describes it — the technical or procedural mechanism that is supposed to make it true.
- **Evidence** — what record exists that the mechanism operated? Design documentation, configuration state, logs, test results, and observed behavior are different kinds of evidence with different strength.
- **Provenance** — where did the evidence come from, and can its origin be trusted? Self-reported evidence from the system being assessed carries different weight than independently observed evidence.
- **Verification** — has the evidence actually been examined, or only collected? A folder of logs that nobody read is not verification.
- **Alternative explanation** — is there another explanation for what was observed, besides the one that supports the desired conclusion? A control that appeared to work might have simply never been tested by anything that would have revealed its absence.
- **Finding** — a specific, evidence-supported statement about what was and was not observed.
- **Limitation** — what the finding does not cover: untested paths, unavailable evidence, time-bound observations, scope exclusions.
- **Bounded conclusion** — a conclusion that stays inside what the evidence actually supports, stated plainly enough that a reader cannot mistake its scope for something broader.

This chain reappears throughout the guide, and every worked example follows a consistent three-part close:

> **Assessment question:** the specific question tested under defined conditions.
>
> **Evidence to preserve:** test or observation identity and version, execution conditions, observed results — including failures and unsuccessful attempts — limitations, and alternative explanations.
>
> **Conclusion boundary:** what the result supports for the tested paths and conditions specifically, and what it does not establish beyond that.

5. Recurring Examination Dimensions

Four broad dimensions recur throughout an AI system examination. They are not a taxonomy, control baseline, maturity model or scoring structure, and they do not replace the applicability analysis developed in Part IV. They are not scored, weighted, or combined into a single number — they are simply where the questions tend to cluster:

- **Security and access control** — resistance to adversarial input, unauthorized access, and unintended lateral movement through connected systems.
- **Data provenance and pipeline integrity** — where data came from, whether it may lawfully be used for its intended purpose, and whether the pipelines that move data through the system are monitored and resistant to tampering.
- **Model and system behavior** — whether the system behaves within its intended and authorized bounds, including how it handles inputs it was not designed for.
- **Operational resilience and monitoring** — whether the infrastructure supporting the system can be observed, whether failures are detected, and whether the system can be constrained or reverted when something goes wrong.

This guide does not rank or weight these dimensions against one another, and its Parts are not organized around them. An organization's own risk context determines which dimension deserves the most attention for a given system — this guide does not.

6. Assurance Language: What a Finding Can and Cannot Say

Assessment findings are frequently misread — sometimes by the people who wrote them. A finding that a control operated correctly during the period and scope examined is not the same statement as a finding that the control will always operate correctly, or that the system is now safe in some general sense.

This guide requires every finding to be stated at the scope the evidence actually supports:

- "The rate-limiting control was observed to function correctly for the 14-day monitoring period examined, across the request volumes observed during that period" is a bounded, defensible finding.
- "The system is protected against denial-of-service" is not a finding — it is a conclusion the evidence has not been shown to support, phrased as though it had been.

Practitioners preparing findings for board, executive, or regulatory audiences will feel pressure to compress the first kind of statement into the second. This guide's position is that the compression is the failure, not the caution.

7. Independence and Conflicts

An assessment's credibility depends partly on whether the person examining the evidence has an interest in a particular conclusion. This guide does not mandate a specific independence model — organizations vary in size, resources, and regulatory context, and a small organization may not have the staff for a fully independent internal assessment function.

What this guide does require is that independence, or its absence, be stated as part of the assessment record. A finding produced by the same team that built and operates the system being assessed is not automatically wrong, but it is a different kind of evidence than a finding produced by a separated internal assurance function or an external practitioner, and the assessment record should say which one occurred.

8. Relationship to Other ODA3 Practitioner Guides

This guide is the examination companion to an existing pair of ODA3 practitioner guides, and it deliberately reuses their vocabulary rather than introducing a competing one:

- **COM-019 — The AI System Owner's Guide to Operational Assurance** describes the system owner's perspective: maintaining an operating basis, tracking evidence, and knowing what changed. This guide describes the examination perspective on largely the same evidence — someone examining what the system owner has maintained, rather than maintaining it themselves. Where COM-019 establishes a distinction (for example, between an evidence-source state and an implementation-responsibility state, or between "Unknown" and "checked and not observed"), this guide uses that same distinction rather than inventing a parallel one.
- **COM-015 — Building an AI Assurance Evidence Pack** owns evidence-pack preparation, indexing, and provenance treatment in detail. This guide references that discipline where relevant and does not duplicate it.
- **COM-017 — Implementing GAISSF™** owns control implementation guidance. This guide assesses against implementation; it does not re-teach how to implement a control.

An assessment conducted using this guide's method can draw on operational records a system owner maintained per COM-019, and its output can feed into an evidence pack prepared per COM-015 — but this guide does not require that either of those guides has been used first. It is written to be usable standalone.

9. What "Notably Absent" Means in This Guide

Consistent with ODA3's broader publication discipline, this guide identifies what it deliberately does not cover, rather than allowing silence to be read as completeness. The concluding Notably Absent section states this fully; the boundary that matters most for Part I is stated here because it governs everything that follows:

This guide does not describe ODA3's own assessment or certification programme, does not grant or imply assessment authority, and does not establish that completing its methodology produces an ODA3-recognized outcome. Detailed functional-safety determination, sector-specific safety approval, and physical return-to-service authority are also outside this guide's scope — where a system can otherwise influence physical action, Part VIII identifies the physical consequence path and hands off detailed treatment to PAI-SF™ and the applicable practitioner guidance, rather than attempting to make a safety determination here.

Part I Checkpoint

Can the reader, after this Part, state in one sentence what an assessment following this guide does and does not establish? If not, Part I has not done its job — the rest of the guide assumes this distinction is already understood.

The checkpoint is an orientation aid, not a completion requirement or readiness gate.

PART II — DEFINE THE ASSESSMENT QUESTION

10. Why the Question Comes Before the System

It is tempting to begin an assessment by describing the system: what it is, what it does, how it is built. Part III does exactly that. But an assessment that starts with system description before the question is defined tends to produce a survey rather than an assessment — a broad description of everything observable, with no way to tell afterward whether anything material was actually tested.

This guide requires the assessment question to be defined first, in writing, before evidence collection begins. The question determines scope; scope determines what evidence is relevant; and only once both are fixed does examining the system become an assessment rather than an inventory.

11. What Makes a Question Assessable

An assessable question has four properties. A question missing any one of them needs to be narrowed before work begins.

- **It names a specific system, version, and period.** "Is our chatbot secure?" names none of these. "Does the customer-support chatbot deployed to production as of the current release, examined over the most recent 30-day operating period, correctly refuse to execute account-modification actions without a second verification step?" names all three.
- **It names a specific claim or concern.** The question should state what would make the answer matter — a claim being tested (a policy asserts X), a concern being investigated (a stakeholder worries about Y), or a requirement being verified (a control baseline requires Z).
- **It has a defined evidence path.** Before starting, it should be plausible to say what kind of evidence — logs, configuration, test results, interview, documentation — could actually answer the question. A question with no imaginable evidence path is not yet assessable; it needs to be reformulated until it has one.
- **It admits a bounded answer.** The question should be answerable in a way that can be qualified — "yes, for the tested conditions," "no, with this exception," "unknown, because evidence was unavailable" — rather than only admitting an unqualified yes or no.

12. The Seven Elements of an Assessment Question Record

Every assessment conducted under this guide should produce a written Assessment Question Record before evidence work begins. The record has seven elements:

1. **System or service identity** — the specific system, by name and identifier, that the question concerns. Where the organization maintains a system inventory or a COM-019-style System Owner Record, reference it directly rather than re-describing the system from scratch.
2. **Claim or concern being examined** — what assertion, worry, or requirement motivates the question. This can be as simple as "a specific control in our applicable baseline requires this" or as open as "a board member asked whether this is true."
3. **Purpose of the assessment** — why this question, now. Common purposes include pre-deployment review, periodic reassessment, response to an incident or near-miss, vendor/third-party due diligence, and preparation for an external review. The purpose affects how the findings will be used, which affects how much rigor and independence the assessment needs.
4. **Consequence if the claim is false** — what happens if the system does not actually do what the claim asserts. This is not a risk score; it is a plain description used to judge how much assessment effort the question deserves.
5. **Period and version** — the specific time window and system version/configuration the assessment covers. A finding about last quarter's model version is not automatically a finding about the version currently deployed.
6. **Environment** — production, staging, a sandboxed replica, or another environment, stated explicitly. Findings from a non-production environment carry a different evidentiary weight and should never be silently presented as though they describe production behavior.

7. **Authority and independence** — who requested the assessment, who is conducting it, and whether the assessing party has an interest in a particular outcome (see Part I, Section 7).

13. Worked Example: From a Vague Concern to an Assessable Question

Starting concern (not yet assessable): "We're worried our AI agent might do something it shouldn't."

This concern fails all four assessability tests: no named system/version/period, no specific claim, no clear evidence path, and no bounded answer — "does something it shouldn't" cannot be answered yes or no in any useful way.

Narrowing pass 1: "We're worried the procurement agent might approve purchase orders it isn't authorized to approve."

Better — now there is a specific claim (the agent should only approve orders within its authorized scope) and a plausible evidence path (approval logs, authorization configuration). Still missing system version, period, and environment.

Narrowing pass 2 — assessable form:

> **Assessment question:** For the procurement agent deployed to production as of [date], examined over the [defined period], what evidence exists that the agent's purchase-order approvals remained within its configured authorization limits (spending threshold, vendor allowlist, approval-chain requirements), and what evidence — if any — shows an approval outside those limits?
>
> **Evidence to preserve:** approval-log entries for the period, the authorization configuration in effect during that period (and any changes to it), any approvals flagged for manual review, and any approvals that could not be matched to a documented authorization rule.
>
> **Conclusion boundary:** a finding here supports a conclusion about approvals that occurred and were logged during the examined period, under the authorization configuration in effect at the time. It does not establish that the same result would hold under a different configuration, a different time period, or for approval paths not captured by the logging mechanism examined — those would need to be separately identified as limitations (Part IX) or additional questions.

This is the level of specificity every assessment question in this guide's method should reach before evidence collection begins.

14. Multiple Questions, One Assessment

A single assessment engagement often needs more than one assessment question — for example, an organization assessing a new AI system before deployment may need separate questions for data handling, access control, output behavior, and operational monitoring. This guide does not require collapsing multiple questions into one; it requires each question to independently meet the assessability test in Section 11, and each to receive its own Assessment Question Record.

Resist the temptation to write one broad question and treat sub-topics as informal detail underneath it — this is exactly how a survey masquerades as an assessment. If a sub-topic deserves its own bounded conclusion, it deserves its own question record.

15. When the Question Changes Mid-Assessment

Assessment work sometimes reveals that the original question was based on a mistaken premise, or that a more important question exists nearby. This is a normal and legitimate outcome — but it should be handled explicitly, not silently.

When the question changes materially during an assessment:

- Preserve the original Assessment Question Record rather than editing it in place.
- Create a new record for the revised question, with a note explaining what prompted the change.
- If work already done under the original question remains relevant, carry it forward explicitly with a note on what it does and does not still support under the revised question.

This preserves the assessment's own history — a reader examining the finished work later should be able to see the actual path the assessment took, not a retroactively tidied version of it.

16. Resource-Planning Prompts for This Part

The following prompts support planning the resources an assessment question will require. They are planning aids, not ODA3-supplied benchmarks — the organization supplies every value:

- Internal roles required: _____
- Estimated practitioner hours by phase: _____
- External specialist support required: _____
- Testing/environment resources required: _____
- Evidence-access dependencies: _____
- Legal/privacy review required: _____
- Expected scheduling constraints: _____
- Assumptions and uncertainties: _____

Part II Checkpoint

Can the reader take a vague organizational concern and produce a written Assessment Question Record meeting all seven elements in Section 12? If not, Part II has not done its job — Parts III onward assume the reader can already do this.

The checkpoint is an orientation aid, not a completion requirement or readiness gate.

PART III — ESTABLISH THE SYSTEM BOUNDARY

17. Why the Boundary Comes After the Question

Part II established what is being asked. Part III establishes what, exactly, the question applies to.

That order matters.

If the system is mapped before the assessment question is defined, the work can easily become an inventory exercise: a broad description of components, integrations and dependencies without a clear basis for deciding which of them matter to the examination.

If the question is defined without a system boundary, the opposite problem appears. A practitioner may reach a finding without a stable referent for the system, configuration, authority path, provider dependency or data flow to which that finding applies.

A system boundary is therefore not simply an architecture diagram. For COM-020, it is an **assessment-scope record**: a defensible description of the components, relationships and exclusions necessary to answer the defined assessment question.

The boundary should allow a practitioner who was not involved in the original scoping discussion to determine:

what is inside the examined system; what is outside it; which versions and configurations matter; which identities and authorities can materially affect the result; which provider-controlled dependencies matter; and which consequence paths connect the AI system to business, human or physical outcomes.

The boundary does not need to describe every component in the enterprise. It needs to describe the components material to the question.

18. What Makes an AI-System Boundary Different

AI systems often introduce or intensify boundary conditions that conventional application inventories do not capture adequately.

The distinction is not that ordinary software lacks providers, identities, external services or data pipelines. Rather, AI systems can make those relationships unusually consequential to system behavior.

A model can change independently of application code. A retrieval corpus can change what information becomes available to the model without a software release. A tool definition or delegated identity can change what an agent is capable of doing. A provider may alter model behavior while the customer's local integration remains unchanged. A human-approval mechanism may be part of the effective execution boundary even though it sits outside the model itself.

For assessment purposes, an AI-system boundary may therefore need to include:

- models and model services;
- orchestration and prompt layers;
- retrieval systems and source corpora;
- agents and delegated execution paths;
- tools, APIs and external services;
- human users and decision authorities;
- workload, service and agent identities;
- provider-controlled dependencies;
- data entering, leaving or influencing the system;
- monitoring and containment surfaces;
- and, where applicable, physical consequence paths.

These are examination dimensions, not a new inventory taxonomy. The exact boundary is determined by the assessment question established in Part II.

19. Identify the Components Material to the Question

The boundary record should identify the system components that could materially affect the answer to the assessment question.

For the **model and orchestration layer**, record the model or model service involved, the relevant release or provider reference where available, any fine-tuning or adaptation layer, system instructions or orchestration logic material to the question, and the environment in which the system operates.

For **retrieval-augmented systems**, identify the retrieval source, index or equivalent retrieval layer, the corpus or source collections in scope, the authorization path governing retrieval, and any material retrieval configuration that could change the answer to the question.

For **agentic systems**, identify the sponsoring principal, agent identity, delegated authority, tools or services available to the agent, approval or escalation paths, and the point at which authority is evaluated before consequential execution.

For **identity and access**, identify the human, service and workload identities that can initiate, modify, approve, constrain or observe the relevant behavior. Where a role or delegated identity is material to the question, preserve the distinction between the identity that sponsored an action and the identity that actually exercised authority at execution.

For **providers and shared services**, identify the provider-controlled or shared components that materially influence the assessed property. A provider need not be inside the organization's administrative control to be inside the assessment boundary.

For **data flows**, identify the data that materially influences the assessed property: user input, retrieved content, contextual memory, operational state, training or fine-tuning data where relevant, and the destinations to which outputs or actions are sent.

For **monitoring and containment**, identify the surfaces through which the relevant behavior can be observed, restricted, revoked, isolated or stopped.

For **physical AI**, identify the path from digital decision to physical consequence where that path is material to the assessment question.

A useful boundary record therefore answers not merely "What components exist?" but:

Which components could cause the answer to this assessment question to change?

20. Preserve Implementation Responsibility and Evidence Source Separately

COM-020 inherits an important distinction from COM-019.

Implementation responsibility asks who controls or materially operates a mechanism.

The public implementation-responsibility states used by the practitioner suite include:

Enterprise-controlled, Provider-controlled, Shared and Visibility-limited.

Evidence-source state asks what evidence the organization can actually obtain or observe about that mechanism.

The evidence-source states are:

Customer-observed, Provider-provided, Shared-interface, Provider-only-inaccessible and Unknown.

These are different questions and must remain separate in the assessment record.

A provider-controlled mechanism may still generate strong customer-observed evidence.

An enterprise-controlled mechanism may still be poorly instrumented and supported by weak or incomplete evidence.

A shared mechanism may produce evidence through a shared interface.

A visibility-limited implementation may leave the organization with provider representations but no independently observable operating evidence.

The assessment boundary should therefore record both dimensions where material:

> **Implementation responsibility:** Who operates or controls this mechanism?
>
> **Evidence-source state:** What evidence about that mechanism is actually available to the organization?

Do not infer evidence sufficiency from implementation responsibility alone.

Part VI examines the evidence basis in detail. Part III only establishes where the relevant evidence is likely to originate and where visibility limitations already exist.

21. Shared Components and Multi-System Boundaries

A single AI system rarely exists in isolation.

Several systems may use the same model provider, identity platform, retrieval service, tool server, policy-enforcement mechanism or monitoring infrastructure. Those shared dependencies can create concentration risk and can also make one item of evidence relevant to more than one assessment question.

That does **not** mean a finding can automatically be inherited from one system to another.

When a shared component is material to multiple systems, the boundary record should identify the shared dependency and document the basis for any later evidence reuse.

Before carrying evidence or a finding across systems, consider whether the following are materially equivalent:

the component version; configuration; integration path; authority context; data context; operating period; monitoring path; and the specific property being examined.

If those conditions differ, evidence from one system may provide useful context without supporting the same conclusion for the other.

The correct rule is therefore:

> **Shared component does not mean shared conclusion.**

Evidence may be reusable. Findings remain bounded to the system, conditions and question they actually support.

21.1 When the Boundary Changes During Examination

The system boundary can change as examination proceeds. A component may prove to be in the request path, a shared dependency may be identified, a supplier service may be replaced, or an authority or consequence path may be found that the boundary record did not show.

When the boundary changes materially:

- preserve the prior boundary record rather than overwriting it;
- issue a revised boundary version identifying what changed, why, and the evidence or source for the change;
- identify which assessment questions, applicability determinations and evidence are affected;
- determine whether the assessment question must be revised under §15; and
- record any unresolved boundary uncertainty that remains.

A later boundary determination does not retroactively change the scope within which earlier examination work was performed. Any conclusion must remain bounded to the boundary, scope and evidence actually examined at the time.

22. Worked Example — Establishing a Boundary

The following example is fictional and exists only to demonstrate the method. Names, identifiers and configurations are illustrative; they are not ODA3 benchmarks or representations of any real provider deployment.

A fictional organization operates **Support Assistant A**, which answers internal support questions using retrieval from an enterprise knowledge repository.

The assessment question is:

> **Assessment question:** For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

The boundary material to that question includes:

Application/orchestration layer: the deployed assistant configuration and the orchestration path that converts a user request into a retrieval request.

Model service: the model service used to generate the final response. The internal operation of the model provider is not itself the property being examined, but the provider interaction remains inside the dependency boundary because model-service behavior forms part of the response path.

Retrieval layer: the retrieval index, source collections, document-permission metadata and retrieval authorization mechanism.

Identity layer: the authenticated user identity, the application/service identity used for retrieval, and any role or delegated permission that can modify retrieval scope.

Data flow: user request → identity context → retrieval request → authorization decision → retrieved content → model context → generated response.

Monitoring path: retrieval logs, authorization-decision records and response records relevant to the defined period.

Out of scope: model training data, unrelated user roles, administrative interfaces not involved in the tested path, and provider internal controls that do not materially determine retrieval authorization for this question.

Implementation-responsibility states: retrieval configuration may be enterprise-controlled; the model service may be provider-controlled; the end-to-end interaction may include shared responsibilities.

Evidence-source states: authorization logs may be customer-observed; provider service characteristics may be provider-provided; some provider-internal evidence may be inaccessible.

> **Evidence to preserve:** the approved boundary record and version; relevant configuration state; identity and role definitions; retrieval authorization configuration; source and index references; logs or observation records for the assessed period; provider representations relied upon; known inaccessible evidence; explicit exclusions; and any boundary changes discovered during examination.
>
> **Conclusion boundary:** the boundary establishes which components and execution paths are included in the examination of retrieval-time authorization. It does not establish that authorization operated effectively, that provider controls are sufficient, that other user roles behave identically, or that the system is secure generally. Those conclusions require the evidence and verification steps developed in later Parts.

The purpose of the boundary record is therefore not to prove the control.

It is to define **where proof would have to come from**.

23. Physical Consequence Paths

Where an AI system can initiate, modify, authorize, inhibit or materially influence physical action, the system boundary must include the physical consequence path relevant to the assessment question.

A useful representation may include:

environment → sensing → data interpretation → inference or planning → authority decision → command → actuator or physical process → safety control → human intervention → physical outcome

Not every physical-AI system will use this exact chain. The boundary record should reflect the actual architecture.

The practitioner should identify:

the physical action the AI can influence; the digital and physical components between inference and consequence; the authority required before action occurs; independent safety or containment mechanisms; human intervention points; and the evidence available across the digital and physical portions of the path.

COM-020 includes examination of this consequence path and the evidence relevant to the assessment question.

It does **not** make a functional-safety determination, grant sector-specific safety approval, or authorize physical return to service.

PAI-SF™ and the applicable practitioner guidance should be used for detailed physical-AI security architecture and control implementation. Functional-safety determination and return-to-service authority remain with the applicable safety regime and authorized decision authority.

This distinction should be preserved explicitly in every physical-AI assessment record.

24. Practitioner Field — System Boundary Record

Use one boundary record for an assessment question or for a group of questions only where the same boundary genuinely applies.

Assessment Question Reference: _____

System / service: _____

Version / release / configuration reference: _____

Environment: _____

Operating period: _____

Components material to the question: _____

Model / model-service layer: _____

Orchestration / system-instruction layer: _____

Agent layer, if applicable: _____

RAG / retrieval layer, if applicable: _____

Tools / external services: _____

Human roles and decision authorities: _____

Human/service/workload identities: _____

Data entering or influencing the assessed behavior: _____

Output/action destinations: _____

Monitoring / telemetry surfaces: _____

Containment / revocation surfaces: _____

Provider / third-party dependencies: _____

Physical consequence path, if applicable: _____

Explicitly out of scope: _____

For each material shared/provider component:

Implementation responsibility: Enterprise-controlled / Provider-controlled / Shared / Visibility-limited / other identified organizational state: _____

Evidence-source state: Customer-observed / Provider-provided / Shared-interface / Provider-only-inaccessible / Unknown: _____

Known boundary uncertainty: _____

Boundary version / date: _____

Boundary change since question definition: _____

Resource-planning prompts may also be recorded for practitioner use:

Internal roles required for boundary mapping: _____

External specialist/provider participation required: _____

Estimated practitioner effort: _____

Evidence-access dependencies: _____

Technical environment/access required: _____

Scheduling constraints: _____

Assumptions and uncertainties: _____

These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration or effort benchmarks.

Part III Checkpoint

Can a practitioner who did not participate in the scoping discussion identify:

what system is being examined; which components can materially affect the assessment question; which versions, identities and provider dependencies matter; which evidence boundaries already exist; which physical consequence paths matter; and what is explicitly excluded?

If not, the assessment boundary is not yet sufficiently defined.

The checkpoint is an orientation aid, not a readiness gate, assessment state or certification prerequisite.

PART IV — DETERMINE APPLICABILITY

25. Why Applicability Must Follow the Boundary

Part II defined the question. Part III defined the system boundary material to that question. Part IV determines which requirements, controls, and regulatory contexts are potentially applicable to that specific system, within that specific boundary, to answer that specific question.

Applicability is not a blanket statement that "System X is subject to Regulation Y." It is a precise mapping exercise. A single AI system may contain components, data flows and authority paths for which different criteria may be relevant. Applicability can therefore differ across the defined system boundary and across assessment questions.

Determining applicability early prevents two common assessment failures:

8. **Over-scoping:** Including criteria that do not bear on the defined question or boundary, increasing examination effort and obscuring material issues.
9. **Under-scoping:** Failing to identify a specific requirement or control that applies to a narrow but critical data flow, identity path, or decision path within the broader system.

The output of this Part is not a legal determination of conformity or a finding about effectiveness. It is a documented, defensible record of which criteria the examination will consider, the authoritative basis for including or not including them, candidate system elements to examine, and explicit exclusions, dependencies, uncertainties and provenance.

26. What "Applicability" Means for an AI Assessment

A defensible finding requires an identified criterion against which the evidence can be evaluated. That criterion may come from a control framework, internal requirement, contract, law or regulation, system specification, or another explicitly identified source.

For COM-020, **applicability** means:

- which criteria are potentially relevant to the frozen assessment question and boundary;
- the authoritative source for each criterion and the applicability rationale;
- which candidate system elements (from the Part III boundary record) appear relevant to each criterion;
- what has been explicitly determined not applicable, with stated basis;
- what remains unresolved and why;
- and where dependencies, shared responsibility, or interpretation uncertainty affect applicability.

Applicability is therefore a **scoping artifact**, not an early assessment engine. It establishes what will be examined in later Parts, not whether controls operate effectively or whether legal conformity exists.

The clean methodological boundary for this Part is:

> **Part IV identifies the criteria potentially applicable to the frozen question and boundary, records the authoritative basis and applicability rationale, identifies candidate system elements to examine, and preserves exclusions, dependencies, uncertainty and provenance. It does not determine control effectiveness, evidence sufficiency, legal conformity, or a finding state.**

27. Identifying the Applicable Control Baseline

An assessment question should be evaluated against explicitly identified criteria. Without an identified criterion, the examination lacks a stable referent.

The applicable baseline may come from one or more of the following sources:

- **Internal organizational policy or standard:** for example, an internal access-control policy requiring retrieval-time authorization equivalent to direct-access authorization.

- **Control framework:** for example, [applicable GAISSF™ control reference] where the framework is in use, or another framework the organization has adopted. Where GAISSF™ is referenced, use a verified GAISSF™ v1.0 reference checked against the published framework, or a deliberately generic placeholder such as `[applicable GAISSF™ control reference]` until the source is verified. Do not present unverified or invented control identifiers as factual GAISSF™ content.
- **Contractual obligation:** for example, a service-level or data-processing requirement.
- **Regulatory or legal requirement used as applicability input:** for example, a data-protection or AI-governance obligation cited as input to examination design (see §30).

Applicability rule: The applicability record must cite the source of each criterion, the jurisdiction, version or effective date where material, and the applicability basis. If a control is derived from a framework, record its identifier and stated intent as published, not as reinterpreted. The examination in later Parts tests the candidate mechanism against the criterion's intent under defined conditions; Part IV does not test effectiveness.

28. Mapping Requirements to Candidate System Elements

A control baseline states what must be achieved. The system boundary (Part III) identifies what exists. This section bridges the two by identifying the candidate mechanism or system element to which the requirement appears relevant.

For each potentially applicable criterion, the applicability record should identify:

10. **The criterion:** what the baseline requires, as stated in its authoritative source.
11. **Candidate system element(s):** the specific element(s) from the Part III boundary record that appear relevant to the criterion — for example, a retrieval authorization mechanism, an identity provider assertion path, or a monitoring surface. This is a scoping identification, not a determination of how the mechanism is intended to satisfy the requirement.

If an applicable requirement has no identified corresponding mechanism at this stage, record **"mechanism not identified at applicability stage"** and carry that condition forward for examination. Do not convert the absence of an identified mechanism into a finding at this stage. Part V examines the intended control mechanism, including intended outcome, implementation mechanism, ownership, operating boundary and failure conditions; Parts VI and VII establish and verify the evidence basis; Part X governs finding formation.

Part IV therefore identifies **where to look**, not whether the control works.

29. Exclusions, Dependencies and Applicability Dispositions

Not every criterion in a broad framework or regulation applies to every system or every assessment question. A well-documented applicability disposition is a sign of rigorous scoping, not assessment avoidance.

Use neutral applicability dispositions rather than categorical permissions:

- **Applicable:** criterion is potentially relevant to the frozen question and boundary, basis stated.
- **Not applicable with stated basis:** criterion does not apply to this question/boundary, basis and authoritative support stated.
- **Addressed under another assessment question:** criterion is material but examined under a separate question record (reference that question).
- **Applicability unresolved:** applicability cannot yet be determined due to missing information, unresolved interpretation, or dependency on further boundary or legal analysis — record what is needed to resolve it.

Critical constraint:

> An exclusion records a scoped applicability position and its basis. It does not establish that a legal, regulatory, contractual or framework obligation is inapplicable unless the authority governing that obligation supports that conclusion.

Dependencies: If a candidate mechanism's operation depends on an external system (for example, an identity provider, a cloud IAM service, a provider-controlled model service), record that dependency, its

implementation-responsibility state and evidence-source state from Part III, and whether the dependency is assumed as a precondition or is itself within examination scope for this question.

Difficulty obtaining evidence, provider control, and resource constraints are not themselves bases for declaring a requirement inapplicable. Record the constraint and carry it into the evidence and verification stages. Its eventual effect on a finding depends on what evidence is available and what the examination establishes.

29.1 When Applicability Changes During Examination

Applicability mapping can change as examination proceeds. A previously unconsidered requirement may become material; an exclusion may no longer be supportable; a regulatory interpretation may be clarified; or a newly discovered boundary element may introduce additional applicable criteria.

When applicability changes materially:

- preserve the prior applicability record rather than overwriting it;
- issue a revised applicability version identifying what changed, why, and the source or authority for the change;
- identify which controls or assessment questions are affected and which evidence is affected;
- determine whether the system boundary or assessment question must be revised under §15 or §21.1; and
- record any unresolved interpretation that remains.

A later applicability determination does not retroactively change the requirements against which earlier examination work was performed. Any conclusion must remain bounded to the criteria, scope and evidence actually examined at the time.

30. Regulatory Context as Applicability Input, Not Determination

AI systems often operate in regulated environments. COM-020 is a methodology guide, not a legal advisory service.

Where legal or regulatory requirements are used as applicability inputs, record:

- the authoritative source, jurisdiction, version or effective date where material;
- the specific requirement as published, not as reinterpreted;
- the person or function responsible for the applicability determination;
- and any unresolved interpretation or dependency on further legal analysis.

COM-020 does not independently determine legal applicability. Regulatory classification and conformity determination are legal and governance functions. The assessment provides evidence of how specific mechanisms operate; legal counsel and governance bodies use that evidence, alongside broader organizational context, to make conformity determinations.

The governing rule is:

> Where legal or regulatory requirements are used as applicability inputs, record the authoritative source, jurisdiction, version or effective date where material, the person or function responsible for the applicability determination, and any unresolved interpretation. COM-020 does not independently determine legal applicability.

The separation to preserve is:

legal/regulatory applicability determination → examination input → mechanism/evidence examination

rather than:

COM-020 methodology → legal classification.

Any worked example referencing regulation must include the disclaimer that the examination tests mechanism operation against stated requirements and does not constitute a legal determination of conformity.

31. Shared Responsibility and Applicability

Shared responsibility affects how an applicable criterion is allocated across the system boundary; it does not by itself determine whether the criterion is applicable. A criterion may remain applicable even where the relevant mechanisms are enterprise-controlled, provider-controlled, shared or visibility-limited. The applicability record should therefore preserve the applicability determination separately from implementation responsibility and evidence-source state.

For each potentially applicable criterion that involves shared or provider-controlled components, the applicability record should preserve:

- **Applicability determination:** Applicable / Not applicable with stated basis / Addressed under another assessment question / Applicability unresolved, with the authoritative basis recorded separately from responsibility allocation.
- **Implementation responsibility per candidate element:** enterprise-controlled / provider-controlled / shared / visibility-limited (from Part III).
- **Evidence-source state per candidate element:** customer-observed / provider-provided / shared-interface / provider-only-inaccessible / unknown (from Part III).
- **Responsibility implication:** which candidate system elements and responsible parties are relevant to examination of the applicable criterion, and where responsibility, visibility or interpretation remains unresolved.

The locked distinction is:

criterion applicability ≠ implementation responsibility ≠ evidence-source state

Shared responsibility does not by itself make a criterion inapplicable, nor does it convert a provider-controlled element into an automatic exclusion. It informs responsibility allocation, the candidate elements that may require examination, and the evidence paths that may be available.

32. Worked Example — Determining Applicability (Illustrative)

> *Illustrative fictional system — system names, internal-policy identifiers, record identifiers and configurations below are invented solely to demonstrate the method. GAISSF™ is a real ODA3 framework; any GAISSF control reference shown as a placeholder must be replaced with a verified reference from the published framework before release. Nothing in this example is legal advice or an ODA3 assessment outcome.*

Assessment question (from Part II): For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

System boundary reference: Part III boundary record — Support Assistant A, retrieval index [knowledge-index-07], support-agent role, defined period.

Applicable criteria (illustrative):

12. Internal policy [SEC-POL-014 §4.2]: "All retrieval systems must enforce document-level access controls at the point of retrieval, equivalent to direct document access." Source: internal policy library, version [X], owner [policy owner role].
13. Control framework criterion: [applicable GAISSF™ control reference — for example, access-enforcement criterion as published in GAISSF™ v1.0]. Source: GAISSF™ v1.0, [control identifier verified against published framework or placeholder pending verification].

Candidate system elements (from Part III): retrieval API gateway, document-level ACL enforcement mechanism in retrieval path, identity context propagation from identity provider, retrieval authorization decision records.

Applicability disposition:

- Internal policy [SEC-POL-014 §4.2] — **Applicable**. Basis: retrieval path handles documents subject to document-level ACLs; assessment question directly concerns retrieval-time enforcement.
- [Applicable GAISSF™ control reference] — **Applicability disposition to be completed after the verified control text and intent are inserted.**

- Training-data provenance criteria — **Addressed under another assessment question**. Reference: separate question record for training-data provenance. Basis: this question is scoped to runtime retrieval authorization.
- Human-oversight criteria — **Not applicable with stated basis / Applicability unresolved** depending on organizational framework interpretation. Basis: human oversight may apply at downstream decision layer; requires clarification from governance function responsible for applicability determination. Recorded as unresolved pending that input.

Mechanism identification at applicability stage:

- For retrieval-time ACL enforcement: candidate mechanism identified — retrieval API gateway ACL enforcement.
- For any criterion where no corresponding mechanism was identified in the current boundary: **"mechanism not identified at applicability stage"** — carried forward to Part V for examination of intended mechanism.

Dependencies: relies on identity provider to correctly assert support-agent role; evidence-source state: provider-provided assertion with customer-observed logs; implementation responsibility: shared (IdP provider-controlled, application enforcement enterprise-controlled).

Exclusion basis (example): Any criterion determined not applicable records the authority supporting that position, not merely practitioner judgment.

> **Assessment question:** For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?
>
> **Evidence to preserve:** applicability record version and date; authoritative sources cited with jurisdiction/version/effective date where material; applicability rationale for each criterion; candidate system elements identified from Part III; exclusions with stated basis and authority; dependencies with implementation-responsibility and evidence-source states; unresolved interpretations; prior applicability versions; boundary and question references; known inaccessible evidence about applicability.
>
> **Conclusion boundary:** This applicability record identifies the criteria potentially applicable to the frozen question and boundary, records authoritative basis and applicability rationale, identifies candidate system elements to examine, and preserves exclusions, dependencies, uncertainty and provenance. It does not determine control effectiveness, evidence sufficiency, legal conformity, or a finding state. Those determinations belong to later Parts and to the governance and legal functions responsible for conformity.

33. Practitioner Field — Applicability Record Template (Blank)

Use one record per assessment question, or per group of questions only where the same criteria genuinely apply.

Assessment Question Reference: _____

System Boundary Reference: _____

Applicability Record Version / Date: _____

Applicable criteria:

- Source (internal policy / framework / contract / regulation as applicability input): _____
- Specific criterion identifier / requirement text as published: _____
- Authoritative source reference, jurisdiction, version/effective date where material: _____
- Person/function responsible for applicability determination: _____
- Applicability disposition: Applicable / Not applicable with stated basis / Addressed under another assessment question / Applicability unresolved: _____
- Basis for disposition and authority supporting it: _____
- Candidate system element(s) from Part III boundary: _____ (or "mechanism not identified at applicability stage")
- Implementation responsibility per candidate element: Enterprise-controlled / Provider-controlled / Shared / Visibility-limited: _____

- Evidence-source state per candidate element: Customer-observed / Provider-provided / Shared-interface / Provider-only-inaccessible / Unknown: _____
- Dependencies and assumptions: _____
- Unresolved interpretation or information needed to resolve applicability: _____

Exclusions / Not applicable:

- Criterion excluded: _____
- Stated basis and authority: _____
- Reference to separate assessment question if applicable: _____

Regulatory context as input:

- Regulatory requirement cited as applicability input: _____
- Source, jurisdiction, version/effective date: _____
- Person/function responsible for legal applicability determination: _____
- Disclaimer: This record maps technical mechanisms to regulatory requirements for examination purposes. It does not constitute a legal determination of compliance or conformity. _____

Prior applicability record version / reference: _____

Material change in applicability discovered during examination and effect on existing evidence: _____

Boundary or question revision required under §15 or §21.1: _____ (yes/no, reference)

Resource-planning prompts (completed by the practitioner):

Internal roles required for applicability mapping: _____

External specialist/legal/governance participation required: _____

Estimated practitioner effort: _____

Evidence-access dependencies (e.g., policy library, framework source, regulatory text): _____

Technical environment/access required: _____

Scheduling constraints: _____

Assumptions and uncertainties: _____

These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration or effort benchmarks.

Part IV Checkpoint

Can a third party read this applicability record and understand exactly which criteria are potentially applicable to the frozen question and boundary, the authoritative basis for each inclusion or exclusion, which candidate system elements are to be examined, which dependencies and uncertainties remain, and that this is a scoping artifact — not a determination of effectiveness, evidence sufficiency, legal conformity, or a finding state?

If not, the applicability mapping is not yet sufficiently defined.

The checkpoint is an orientation aid, not a readiness gate, assessment state or certification prerequisite.

PART V — UNDERSTAND THE CONTROL

34. Why Understanding the Control Comes After Applicability

Part II defined the assessment question. Part III defined the system boundary. Part IV identified which criteria are potentially applicable and which candidate system elements appear relevant to each criterion.

Part V now examines the **represented intended control model** for each candidate system element — the intended outcome, represented mechanism, ownership and responsibility, intended operating boundary, dependencies, and hypothesized failure conditions.

This order matters because Part V must not pre-decide what Part IV has not yet established. If Part V begins by documenting how a mechanism operates before Part IV has identified whether the criterion even applies, the examination has already jumped ahead of its own scoping.

Part V establishes the **represented intended state** of the control. It does not establish whether the control **actually operates** that way.

Part VI builds the evidence basis by identifying, collecting or preserving relevant evidence and its provenance; Part VII verifies relevant propositions against that evidence; Part X governs finding formation. Part V does not perform any of those later functions.

The clean methodological boundary for this Part is:

Part V documents the represented intended control model for each candidate system element identified in Part IV: intended outcome, represented mechanism, ownership and responsibility, intended operating boundary, dependencies, and hypothesized failure conditions.

It also preserves the provenance and limitations of that representation. It does not establish that the mechanism exists or is configured as represented, collect evidence of actual operation, determine evidence sufficiency, verify whether the mechanism works, establish that a failure occurred, or produce a finding state. Those determinations belong to later Parts.

35. What "Understanding the Control" Means

For each candidate system element identified in Part IV, Part V documents six dimensions of the **represented intended control model**:

14. Intended outcome: What the control is supposed to achieve, stated in terms that can later be examined.
15. Represented mechanism: How available design, configuration, procedural, provider, or owner representations say the candidate system element is intended to implement the criterion.
16. Ownership and responsibility: Which roles or parties are represented as operating, maintaining, governing, or being accountable for the mechanism.
17. Intended operating boundary: The conditions and scope within which the mechanism is represented or intended to function.
18. Dependencies: The systems, identities, data, configurations, interfaces, procedures, or other preconditions on which the represented mechanism depends.
19. Hypothesized failure conditions: States or behaviors that would be inconsistent with the represented intended operation and can later inform evidence collection and verification.

These six dimensions form the **Control Understanding Record** for each candidate system element.

The record must also preserve the source and provenance of the representation used to describe each dimension — for example, a design document, configuration specification, architecture record, provider documentation, policy, procedure, system-owner statement, or another identified source, with version or date where material.

Recording a source here does not establish that the representation is accurate, current, complete, implemented, or sufficient as evidence.

Part V therefore establishes a documented model of what the control is represented or intended to do. It does not establish that the mechanism exists or is configured as represented, that it operated during the relevant period, that available evidence is sufficient, or that any finding state applies.

36. Identify the Intended Outcome

The intended outcome is what the control is supposed to achieve, stated in observable terms.

For each candidate system element identified in Part IV, the Control Understanding Record must state:

- The criterion being satisfied: The specific requirement or control from the applicable baseline (e.g., "Internal Policy SEC-POL-014 §4.2 requires retrieval-time authorization equivalent to direct document access").
- The intended outcome in examinable terms: What behavior or state would be consistent with the control's stated intent and could later be examined (e.g., "When a user authenticated as support-agent requests retrieval of a document, the retrieval path enforces the same document-level access controls that apply to direct document access for that user").

The intended outcome should be stated precisely enough that later evidence collection and verification can examine it. A statement like "the system is secure" is not an intended outcome. A statement like "the retrieval path enforces document-level access controls at retrieval time equivalent to direct document access enforcement" is an intended outcome.

The intended outcome should remain traceable to the criterion's authoritative source and any identified organizational interpretation.

Part V should distinguish published criterion text from practitioner interpretation rather than silently rewriting the criterion.

37. Identify the Represented Mechanism

The represented mechanism describes how identified sources say the candidate system element is intended to implement the criterion. It is a control-design representation, not an established fact about implementation or operation.

For each candidate system element, the Control Understanding Record should identify:

- Represented technical or procedural mechanism: The component, configuration, process, procedure, or combination that is represented as achieving the intended outcome.
- Represented implementation location: Where within the Part III system boundary the mechanism is represented as residing or operating.
- Representation source and provenance: The design document, configuration specification, architecture record, provider documentation, policy, procedure, system-owner statement, or other source from which the description is derived, including version or date where material.
- Representation limitations or unresolved points: Any material ambiguity, conflict, incompleteness, or unverified assumption in the description.

The represented mechanism is not the same as the intended outcome. The intended outcome describes what is supposed to be achieved; the represented mechanism describes how identified sources say that outcome is intended to be achieved.

A single intended outcome may be represented as relying on multiple mechanisms, and a single mechanism may contribute to multiple intended outcomes. Part V records those relationships without determining whether the mechanisms exist, are configured as represented, or operate effectively.

38. Ownership and Responsibility

Ownership and responsibility record the roles or parties represented as operating, maintaining, governing, or being accountable for the mechanism. They do not establish examination authority, evidentiary strength, or control effectiveness.

For each candidate system element, the Control Understanding Record should identify, where applicable:

- Operational owner: The role or party represented as responsible for day-to-day operation.

- Maintenance owner: The role or party represented as responsible for maintenance or configuration.
- Accountable organizational function: The function represented as accountable for the control or its governance.
- Relevant decision or escalation authority: Any practitioner-identified authority relevant to material control decisions or unresolved issues.
- Implementation-responsibility state: Enterprise-controlled / Provider-controlled / Shared / Visibility-limited, carried forward from Part III.
- Evidence-source state: Customer-observed / Provider-provided / Shared-interface / Provider-only-inaccessible / Unknown, carried forward separately from Part III.

The locked distinction remains:

ownership/accountability ≠ implementation responsibility ≠ evidence-source state ≠ examination authority Do not infer evidentiary strength from ownership or implementation responsibility. Part V records the responsibility context and the source of the representation; later Parts determine what evidence is available and what it can support.

39. Intended Operating Boundary

The intended operating boundary records the conditions and scope within which the mechanism is **represented or intended** to function. It is not a verified operating envelope.

For each candidate system element, the Control Understanding Record should identify:

- Represented intended scope: What the mechanism is represented as covering.
- Represented intended conditions: The conditions under which the mechanism is represented as expected to operate.
- Known or represented limitations: Documented limitations, exclusions, assumptions, or conditions outside the represented design.
- Relationship to the assessment question: Whether any represented limitation or condition is material to the frozen assessment question.

The intended operating boundary is not the same as the assessment question's scope from Part II. The assessment question defines what is being examined; the intended operating boundary records what identified sources say the mechanism is designed or intended to cover.

If the represented operating boundary is narrower than the assessment question, Part V records that relationship and carries it forward. It does not determine at this stage whether the limitation is acceptable, whether another mechanism is required, or whether a finding follows.

40. Hypothesized Failure Conditions

Failure conditions are hypotheses about states or behaviors that would be inconsistent with the represented intended operation. Recording them helps later evidence collection and verification define what to examine.

It does not establish that a failure occurred or determine a finding state.

For each candidate system element, the Control Understanding Record should identify, where relevant:

- Observable inconsistent states or behaviors: Outcomes that would conflict with the represented intended operation.
- Hypothesized failure modes: Ways the represented mechanism could fail, be bypassed, become stale, be misconfigured, or operate through an alternate path.
- Dependency failures: Failure or degradation of an identity, data, provider, interface, configuration, procedural, or other dependency on which the mechanism relies.
- Silent or incomplete failure: Conditions in which the mechanism may fail without an expected log, alert, refusal, or other visible signal.
- Represented detection or escalation path: How identified sources say such conditions are intended to be detected, recorded, refused, escalated, or otherwise handled.

These are examination hypotheses, not findings and not proof that the listed failure modes are possible in the actual implementation. Part VI may use them to identify, collect or preserve evidence relevant to the assessment question and to document that evidence's provenance. Evidence collection or

preservation in Part VI does not by itself establish sufficiency or that the represented mechanism operated as intended. Part VII verifies relevant propositions against the evidence. Part X governs finding formation.

40.1 When the Control Understanding Changes During Examination

The represented control model may change as examination proceeds. A design document may be superseded, a system owner may correct an earlier description, a configuration may reveal a different intended mechanism, a dependency may become material, or conflicting representations may emerge.

When the Control Understanding Record changes materially:

- preserve the prior record rather than overwriting it; issue a revised version identifying what changed, why, and the source of the change;
- identify which intended outcomes, represented mechanisms, responsibility assignments, operating-boundary statements, dependencies, or failure hypotheses are affected;
- identify any evidence already collected against the earlier representation; determine whether the assessment question, system boundary, or applicability record requires revision under §15, §21.1, or §29.1;

and preserve unresolved conflicts between representations rather than silently selecting one.

A later Control Understanding Record does not retroactively redefine what earlier examination work addressed. Any later conclusion must remain bounded to the control representation, system state, scope, period, and evidence actually examined.

41. Worked Example — Understanding a Control (Illustrative)

Note: *Illustrative fictional system — system names, internal-policy identifiers, record identifiers, role names, design descriptions, configurations, dependencies, and operating assumptions below are invented solely to demonstrate the method. They are represented-design facts for this fictional example, not observations about a real system and not ODA3 assessment evidence. GAISSF™ is a real ODA3 framework; any GAISSF control reference shown as a placeholder must be replaced with a verified reference from the published framework before release.*

Assessment question (from Part II): For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Candidate system element (from Part IV): Retrieval authorization element in the fictional retrieval path.

Control Understanding Record — fictional represented design:

Intended outcome:

- **Criterion:** fictional internal policy [SEC-POL-014 §4.2] — represented for this example as requiring retrieval-time authorization aligned with direct document-access authorization.
- **Intended outcome in examinable terms:** the fictional design represents the retrieval path as denying retrieval where the requesting identity would not be authorized for direct access to the requested document.
- **Representation source:** fictional internal policy and fictional architecture record [ARCH-RET-01], versions [X] and [Y].

Represented mechanism:

- **Represented technical mechanism:** fictional architecture documentation describes a retrieval authorization component that receives identity context and document authorization attributes and is intended to make an authorization decision before content is returned to the downstream model.
- Represented implementation location: fictional retrieval path identified in the Part III boundary record. - Representation limitations: whether the deployed configuration matches this representation has not been established in Part V.

Ownership and responsibility:

- **Operational owner:** [fictional practitioner-supplied role].

- **Maintenance owner:** [fictional practitioner-supplied role].
- **Accountable organizational function:**

[fictional practitioner-supplied function].

- **Relevant decision or escalation authority:** [fictional practitioner-supplied authority].
- **Implementation-responsibility state:** [Enterprise-controlled / Provider-controlled / Shared / Visibility-limited — carried from the fictional Part III record].
- **Evidence-source state:**

[Customer-observed / Provider-provided / Shared-interface / Provider-only-inaccessible / Unknown — carried separately from the fictional Part III record].

Intended operating boundary: - Represented intended scope:

[fictional scope to be documented]. - Represented intended conditions:

[fictional authentication, metadata, interface, or configuration conditions to be documented].

- **Behavior for documents without authorization metadata:** [represented behavior to be documented].
- **Behavior for unauthenticated or alternate retrieval paths:** [represented behavior to be documented].
- **Relationship to assessment question:** any condition that could permit content retrieval outside the represented authorization path is material to the question and should be carried forward for later examination.

Dependencies:

- **Identity context:** [fictional dependency and represented source].
- **Document authorization attributes:** [fictional dependency and represented source].
- **Configuration or policy synchronization:** [fictional dependency and represented source].
- **Provider or shared interfaces:** [fictional dependency, responsibility state, and evidence-source state].
- **Unresolved dependency assumptions:**

[to be recorded].

Hypothesized failure conditions:

- authorization decision omitted, bypassed, or applied after content retrieval; stale or incorrect identity or authorization context;
- stale, missing, or inconsistent document authorization attributes; alternate retrieval path that does not invoke the represented authorization mechanism;
- dependency failure that changes authorization behavior; silent failure in which expected authorization records, refusals, or alerts are absent;
- **represented detection or escalation path:** [fictional design representation to be documented].

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence to preserve: Control Understanding Record version and date; authoritative criterion source; sources and provenance of the represented design; intended outcome; represented mechanism and location; ownership/responsibility context;

implementation-responsibility and evidence-source states; intended operating boundary; dependencies; hypothesized failure conditions;

representation limitations and unresolved conflicts; prior Control Understanding Record versions; references to the frozen question, boundary, and applicability record.

Conclusion boundary: This Control Understanding Record documents the represented intended control model for the fictional retrieval-authorization element. It does not establish that the mechanism exists or is configured as represented, that it operated during the relevant period, that the identified sources are sufficient evidence of actual operation, that any hypothesized failure occurred, or that a finding state applies. Part VI builds and preserves the evidence basis and provenance; Part VII performs verification against relevant evidence; Part X governs finding formation. No later-stage conclusion is made in Part V.

42. Practitioner Field — Control Understanding Record Template (Blank)

Use one Control Understanding Record for each candidate system element identified in Part IV, unless multiple elements genuinely share the same represented control model and that basis is documented.

Assessment Question Reference: _____\ System Boundary Reference: _____\ Applicability Record Reference: _____\ Candidate System Element (from Part IV): _____\ Control Understanding Record Version / Date: _____\ Prior Control Understanding Record Version / Reference: _____

Criterion and intended outcome: - Criterion source and identifier:

_____ - Criterion text or authoritative intent reference:

_____

- **Intended outcome in examinable terms:** _____
- **Organizational interpretation, if any, and responsible source/function:**

_____ **Represented mechanism:**

- **Represented technical or procedural mechanism:** _____
- **Represented implementation location (from Part III boundary):** _____
- **Representation source(s) and provenance:**

_____

- **Source version/date where material:** _____
- **Representation limitations, conflicts or unresolved points:** _____ Ownership and responsibility:
- **Operational owner:** _____
- **Maintenance owner:** _____
- **Accountable organizational function:**

_____

- **Relevant decision or escalation authority:** _____
- Implementation-responsibility state (from Part III): Enterprise-controlled / Provider-controlled / Shared / Visibility-limited: _____
- **Evidence-source state (from Part III):** Customer-observed / Provider-provided / Shared-interface / Provider-only-inaccessible / Unknown: _____ Intended operating boundary:
- **Represented intended scope:**

_____

- **Represented intended conditions:** _____
- **Known or represented limitations:** _____
- **Relationship of limitations/conditions to the assessment question:** _____ Dependencies:
- **Identity/authority dependencies:** _____
- **Data/metadata dependencies:** _____
- **Configuration/policy dependencies:** _____
- **Provider/shared-interface dependencies:**

_____

- **Other material preconditions:** _____
- **Unresolved dependency assumptions:** _____ Hypothesized failure conditions:
- **Observable inconsistent states or behaviors:** _____
- **Bypass or alternate-path hypotheses:**

_____

- **Stale/misconfigured state hypotheses:** _____
- **Dependency-failure hypotheses:** _____
- **Silent/incomplete-failure hypotheses:** _____
- **Represented detection/refusal/escalation path:** _____ Control-understanding change provenance:
- **Material change discovered during examination:** _____
- **Source/reason for change:**

_____

- **Dimensions affected: _____**
- **Effect on evidence already collected: _____**
- **Question/boundary/applicability revision required under §15 / §21.1 / §29.1: _____**
- **Unresolved conflicting representations: _____** Resource-planning prompts (completed by the practitioner):
- **Internal roles required for control understanding:**

_____

- **External specialist support required: _____**
- **Estimated practitioner effort: _____**
- Access dependencies for understanding the represented design (e.g., design documentation, configuration specifications, interviews): _____
- **Technical environment/access required: _____**
- **Scheduling constraints:**

_____ - Assumptions and uncertainties: _____ These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration or effort benchmarks.

Part V Checkpoint

Can a third party who was not involved in the control-understanding work read the Control Understanding Record and distinguish:

- what the criterion requires; what outcome is intended; how identified sources represent the mechanism; where that representation came from and what remains uncertain;
- who is represented as responsible; the separate implementation-responsibility and evidence-source states; the intended operating boundary and dependencies;
- the hypothesized failure conditions; and what has not yet been established about actual implementation or operation?

If not, the control understanding is not yet sufficiently defined for later evidence work.

The checkpoint is an orientation aid, not a readiness gate, assessment state or certification prerequisite.

PART VI — BUILD THE EVIDENCE BASIS

43. Why the Evidence Basis Comes After Control Understanding

Part II defined the assessment question. Part III defined the system boundary. Part IV identified the criteria potentially applicable and the candidate system elements. Part V documented the represented intended control model for each candidate system element — intended outcome, represented mechanism, ownership and responsibility, intended operating boundary, dependencies, and hypothesized failure conditions.

Part VI now identifies, collects, preserves, and documents the evidence that could bear on that represented model.

This order matters. Evidence collection without a defined represented control model is directionless — a pile of logs, configurations, and documents with no principled basis for deciding which of them matter.

Conversely, a represented control model without an evidence basis is an untested assertion — a description of what the mechanism is supposed to do, with no record of what actually happened.

Part VI develops the **evidence basis**: the documented body of evidence, its provenance, its classification, its temporal scope, and its known gaps. It does not verify propositions against that evidence, determine whether the evidence is sufficient to support a finding, or produce a finding state. Those determinations belong to Part VII (verification) and Part X (finding formation).

The clean methodological boundary for this Part is:

Part VI identifies, obtains or preserves evidence potentially relevant to the assessment question and represented control model. For each evidence item, it records classification, producer and source relationship, provenance, temporal scope, preservation history, relevance relationship, coverage, access limitations, and known uncertainties. Part VI does not assign automatic evidentiary weight, determine what propositions the evidence ultimately supports, establish evidence sufficiency, verify control operation, determine materiality of an evidence limitation, or form a finding. Part VII evaluates what propositions the evidence supports and the limitations of verification; Part X forms findings from the verified results within the defined conclusion boundary.

44. What "Building the Evidence Basis" Means

For each candidate system element whose represented control model was documented in Part V, Part VI develops the evidence basis across six dimensions:

20. Evidence identification: What evidence exists, or could reasonably be obtained, that bears on the represented intended outcome, mechanism, operating boundary, dependencies, or hypothesized failure conditions?
21. Evidence collection or preservation: How is the evidence obtained, captured, or preserved in a form that can later be examined?
22. Provenance: Where did the evidence come from, who produced it, when, and through what chain?
23. Classification: Is the evidence design evidence, implementation evidence, or operating evidence?
24. Temporal scope and relationship: What period, version, configuration state, or reference state does the evidence represent, and how does that relate to the assessment period defined in Part II?
25. Coverage, missing evidence, and access limitations: What evidence has been identified, what is missing, what is inaccessible, and what remains unknown?

These six dimensions form the **Evidence Basis Record** for each candidate system element.

The record must carry forward relevant system/component context from Part III, while assigning an evidence-source state to each evidence item independently (customer-observed / provider-provided / shared-interface / provider-only-inaccessible / unknown). It must also preserve the relevant implementation-responsibility state (enterprise-controlled / provider-controlled / shared / visibility-limited). The evidence producer, evidence-source relationship, implementation responsibility, and observation or independence context are separate attributes. They inform provenance and later

verification strategy; they do not by themselves assign evidentiary weight, determine what proposition the evidence supports, establish sufficiency, or determine finding state.

Central rule for Part VI:

Part VI develops the evidence basis. It does not verify propositions against the evidence, determine whether the evidence is sufficient for a finding, establish that the represented mechanism operated as intended, or produce a finding state.

The following separations must be preserved throughout Part VI:

- evidence existence ≠ evidence sufficiency — evidence may exist without being sufficient to support a particular conclusion evidence collection ≠ evidence verification — collecting

evidence is not the same as testing propositions against it evidence provenance ≠ evidentiary strength — provenance must be documented but does not by itself determine weight

- **design evidence ≠ implementation evidence ≠ operating evidence**

— these are different kinds of evidence with different evidentiary character coverage description ≠ sufficiency determination — describing what has been identified, is missing, or is inaccessible is not the same as judging whether enough exists

- known gap ≠ finding of Gap — a gap in available evidence is not the same as a finding state

45. Evidence Classification: Design, Implementation, Operating

Evidence about a control falls into three categories. Each category answers a different question and carries different evidentiary character. Part VI must classify each item of evidence accordingly.

Design evidence documents what the mechanism was supposed to do.

Examples include:

- Design specifications, architecture records, and system cards
- Policies, procedures, and standards that describe intended behavior Provider documentation describing represented functionality
- Threat models, data-flow diagrams, and authorization matrices
- Training materials describing intended operation Design evidence establishes represented intent. It does not establish that the represented design was implemented, correctly configured, or operated as intended during the assessment period.

Implementation evidence documents what was actually configured, deployed, or put in place. Examples include:

- Configuration exports, infrastructure-as-code, and deployment records
- Access-control lists, role definitions, and permission matrices as configured
- Model versions, prompt versions, and retrieval-index snapshots as deployed
- Tool allowlists, agent-authority configurations, and human-oversight settings as configured
- Change records showing when configurations were modified Implementation evidence establishes what was in place at a point in time. It does not establish that the implementation operated correctly during the assessment period, that it was not bypassed, or that it produced the intended outcomes.

Operating evidence documents what actually happened during the assessment period. Examples include:

- Logs of actual system behavior (retrieval decisions, authorization decisions, tool invocations, model outputs)
- Monitoring records, alert records, and incident records
- Test results from controlled exercises conducted during the period
- Observed behavior from direct examination (e.g., paired tests, adversarial tests)
- Human-oversight records showing actual decisions made during the period Operating evidence may bear on propositions about events or behavior during a defined period. It is bounded to the period, configuration, conditions, and observation path under which it was produced. Operating evidence from one period does not automatically establish behavior in another period.

Evidence classification describes the kind of proposition an item may bear on; it does not assign an automatic hierarchy of evidentiary strength. Design evidence may support propositions about

represented intent, implementation evidence may support propositions about configured or deployed state, and operating evidence may support propositions about events or behavior during a defined period. What any item actually supports depends on its provenance, temporal relationship, integrity, relevance, coverage, limitations, and the proposition being verified in Part VII.

The classification rule:

Each item of evidence in the Evidence Basis Record must be classified as design, implementation, or operating evidence. Where evidence spans multiple categories (for example, a configuration record that also describes intended behavior), the primary classification must be stated along with any secondary character.

A single control may require evidence from all three categories to be fully examined. Part VI classifies and preserves the evidence; Part VII verifies relevant propositions against it.

46. Evidence Sources and Provenance

Every item of evidence has a source and a provenance. Part VI must document both.

Evidence-source state is assigned to each evidence item. Relevant system/component context from Part III may inform the record, but it does not pre-classify an individual evidence item:

- **Customer-observed:** Evidence directly observed or collected by the organization in its own environment (for example, logs from the organization's retrieval API gateway, configuration exports from the organization's infrastructure).
- **Provider-provided:** Evidence supplied by a provider about the provider's own systems or services (for example, provider documentation, provider-supplied audit reports, provider-supplied model cards).
- **Shared-interface:** Evidence produced through a shared interface between the organization and a provider (for example, shared logging, shared monitoring dashboards, shared audit trails).
- **Provider-only-inaccessible:** Evidence that exists within the provider's environment but is not accessible to the organization (for example, provider-internal training data, provider-internal security logs).
- **Unknown:** The source or availability of the evidence cannot be determined.

Provenance extends beyond the source state to document:

- Who produced the evidence (specific role, team, provider function, or system)
- When the evidence was produced (date, time, version, configuration reference)
- How the evidence was produced (automated log, manual export, interview, observation, test execution)
- Chain of custody (how the evidence moved from producer to the assessment record, including any transformations, exports, or copies)
- Any known limitations on the evidence's reliability (for example, logs that may be incomplete, configurations that may not reflect the full state, interviews that represent a single perspective)

The provenance rule:

Every item of evidence in the Evidence Basis Record must have its source state and provenance documented. Evidence without documented provenance is recorded as such and its evidentiary character is noted as limited.

Provenance does not by itself determine evidentiary weight — that is a matter for Part VII's verification and Part X's finding formation. But missing, incomplete, or unresolved provenance must itself be preserved as a limitation in the record. Whether that limitation is material to a proposition being verified is determined later.

Source-relationship rule:

Evidence producer, evidence-source state, implementation responsibility, and observation or independence context must be recorded separately where relevant. Provider-provided evidence is not automatically weaker, and customer-observed or independently observed evidence is not automatically stronger. These attributes shape provenance and verification strategy; they do not assign evidentiary weight by themselves.

47. Temporal Scope, Relationship, and Currency

Evidence has a time dimension. Part VI records when each item was produced, the state or period it represents, and its relationship to the assessment period defined in Part II. It does not determine in Part VI what propositions that timing ultimately supports.

Temporal scope and relationship may include:

- **Point-in-time evidence:** Evidence representing a specific moment, such as a configuration snapshot or deployment record. Record the timestamp and the state represented.
- **Period evidence:** Evidence generated across a stated span of time, such as logs or monitoring records. Record the start and end of the covered period.
- **Superseded evidence:** Evidence replaced by a later version or state. Preserve the earlier item and record the successor reference and relevant change history.
- **Reference evidence:** Evidence not generated as a record of a particular operating period, such as a policy, specification, design record, or mathematical/technical description. Record its version, issue/effective date where applicable, status, and any known supersession or withdrawal.

The temporal-scope rule:

Every item in the Evidence Basis Record must state its known temporal scope and its relationship to the assessment period. Where the relationship is unclear, outside the assessment period, superseded, or dependent on an unverified version/configuration relationship, record that limitation without deciding in Part VI what proposition the item ultimately supports.

Evidence from outside the assessment period may remain relevant context.

Part VII determines what propositions, if any, can be verified using it.

48. Coverage, Missing Evidence, and Inaccessible Evidence

The evidence basis will rarely be complete. Part VI must document what is available, what is missing, and what is inaccessible — without converting any of these into a finding state.

Coverage describes, as an inventory matter, what evidence has been identified relative to the represented intended control model from Part V. For each dimension of the represented model (intended outcome, mechanism, operating boundary, dependencies, hypothesized failure conditions), the Evidence Basis Record should note whether evidence has been identified, is partially represented in the current basis, has not been identified, or remains inaccessible. These are descriptive coverage states, not judgments about adequacy or sufficiency.

Known gaps are specific areas where evidence that would be relevant to the assessment question is not available. Examples include:

- No operating logs exist for a particular component during the assessment period
- Configuration records are incomplete for a particular subsystem
- Provider documentation does not address a particular represented behavior No test was conducted for a particular hypothesized failure condition Known gaps must be documented with:
- What evidence is missing
- Why it is missing (for example, not collected, not retained, not produced by the system, not accessible)
- What assessment question or represented-model dimension it relates to
- Whether the gap can be closed by later evidence collection or is permanent Inaccessible evidence is evidence that exists but cannot be obtained by the organization. Examples include:
- Provider-internal training data (provider-only-inaccessible)
- Evidence protected by legal privilege or regulatory restriction
- Evidence in a system to which the organization has no access
- Evidence that would require provider cooperation that has not been granted Inaccessible evidence must be documented with:
- What the evidence is (to the extent it can be described)
- Why it is inaccessible
- What evidence-source state applies (typically provider-only-inaccessible or unknown)
- What assessment question or represented-model dimension it relates to

- Whether alternative evidence exists that could partially address the same question The coverage rule:

Part VI records what evidence has been identified, what is missing, and what is inaccessible. It does not determine whether the available evidence is sufficient to support a proposition or finding, whether a missing or inaccessible item is material to verification, or whether any finding state is appropriate. Those determinations belong to Part VII (verification) and Part X (finding formation).

A known gap is a descriptive fact about the evidence basis. It is not a finding. A finding about what the gap means — or whether the available evidence is sufficient despite the gap — belongs to later Parts.

48.1 When the Evidence Basis Changes During Examination

The evidence basis may change as examination proceeds. New evidence may surface; previously inaccessible evidence may become available; evidence may be found to be corrupted, incomplete, or unreliable; or a hypothesized failure condition from Part V may lead to the discovery of new evidence categories.

When the Evidence Basis Record changes materially:

- preserve the prior record rather than overwriting it; issue a revised version identifying what changed, why, and the source of the change;
- identify which represented-model dimensions, assessment questions, or hypothesized failure conditions are affected;
- identify any verification work already performed against the earlier evidence basis; determine whether the assessment question, system boundary, applicability record, or control-understanding record requires revision under §15, §21.1, §29.1, or §40.1; and preserve any evidence that has been superseded or found unreliable

rather than silently discarding it.

A later Evidence Basis Record does not retroactively alter what evidence was available to, selected for, or used in earlier verification work.

Any later conclusion must remain bounded to the evidence basis, system state, scope, period, and verification actually performed.

49. Evidence Preservation and Integrity

Evidence should be preserved in a manner appropriate to the artifact type, transformation risk, access constraints, and later examination need. Part VI documents the preservation method actually used; it does not prescribe one universal preservation control for every evidence item.

Possible preservation methods include:

- content hashes or other integrity references for suitable digital artifacts; point-in-time snapshots or structured exports;
- retained source records with version and metadata; observation records with date, observer, conditions, and method;
- provider-supplied records preserved as received; and other organization-approved methods appropriate to the evidence type.

For each item, record where applicable:

- how and when it was preserved; the original format or source reference; transformations applied, such as filtering, aggregation, normalization, redaction, or anonymization;
- whether the original remains available; chain-of-custody or transfer information where relevant; and known preservation or integrity limitations.

The preservation rule:

The Evidence Basis Record must document the preservation method and known transformations or integrity limitations relevant to later examination. The appropriate method depends on the evidence item and examination context. Part VI does not require a particular hashing algorithm, immutable-storage technology, signature method, or custody procedure unless such a requirement comes from an identified governing source or organizational methodology.

Preservation does not equal verification. A well-preserved item may still be irrelevant to a proposition, temporally mismatched, incomplete, unreliable, or otherwise limited. Part VII evaluates what propositions the preserved evidence supports and the limitations of that verification.

50. Worked Example — Building the Evidence Basis (Illustrative)

Note: *Illustrative fictional system — system names, internal-policy identifiers, record identifiers, and evidence descriptions below are invented solely to demonstrate the method. They are not ODA3 benchmarks or representations of any real system, provider, or assessment outcome.*

Assessment question (from Part II): For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Represented intended control model (from Part V): Retrieval authorization element in the fictional retrieval path; intended outcome:

retrieval path denies retrieval where requesting identity would not be authorized for direct access; represented mechanism: retrieval authorization component making authorization decision before content returned to downstream model; dependencies include identity context, document authorization attributes, configuration synchronization;

hypothesized failure conditions include authorization decision omitted, stale identity context, alternate retrieval path.

Evidence Basis Record — fictional represented evidence:

Evidence identification:

- Design evidence:
- Fictional architecture record [ARCH-RET-01] describing retrieval authorization component (source: fictional architecture team; evidence-source state: customer-observed;

implementation responsibility: enterprise-controlled)

- Fictional internal policy [SEC-POL-014 §4.2] stating retrieval-time authorization requirement (source: fictional policy office; evidence-source state: customer-observed;

implementation responsibility: enterprise-controlled)

- Fictional provider documentation for retrieval API gateway describing represented authorization functionality (source: fictional provider; evidence-source state: provider-provided;

implementation responsibility: provider-controlled)

- Implementation evidence:
- Fictional configuration export of retrieval API gateway ACL enforcement settings, snapshot taken [date] (source: fictional platform engineering team; evidence-source state:

customer-observed; implementation responsibility: enterprise-controlled)

- Fictional identity provider role-assertion configuration for support-agent role, snapshot taken [date] (source: fictional identity team; evidence-source state: customer-observed;

implementation responsibility: shared — IdP provider-controlled, role definition enterprise-controlled)

- Fictional document authorization metadata schema definition, version [X] (source: fictional document management team; evidence-source state: customer-observed; implementation responsibility: enterprise-controlled)
- Operating evidence:
- Fictional retrieval authorization decision logs covering assessment period [period] (source: fictional retrieval service; evidence-source state: customer-observed;

implementation responsibility: enterprise-controlled)

- Fictional paired-test results from a practitioner-defined set of document/user combinations selected under a documented basis, conducted [date] (source: fictional examination team;

evidence-source state: customer-observed; implementation responsibility: enterprise-controlled)

- Fictional monitoring records for retrieval authorization component covering assessment period (source: fictional monitoring service; evidence-source state: customer-observed;

implementation responsibility: enterprise-controlled)

Evidence provenance:

- Architecture record: produced by fictional architecture team, version [X], date [Y], preserved as PDF with [recorded preservation/integrity reference, if used]
- Configuration export: produced by fictional platform engineering team via automated export tool, snapshot [date/time], preserved as JSON with [recorded preservation/integrity reference, if used]
- Retrieval authorization logs: produced by fictional retrieval service automatically, covering [period], preserved as structured export with [recorded preservation/integrity reference, if used], retention reference [policy/reference]
- Paired-test results: produced by fictional examination team during controlled test exercise [date], preserved as test report with [recorded preservation/integrity reference, if used], test methodology documented in [test-plan reference]

Temporal scope:

- Architecture record: reference evidence, version [X], issue/recording date [Y], supersession status [status]
- Configuration export: point-in-time, snapshot [date/time] — valid for that moment; does not establish configuration at other times
- Retrieval authorization logs: period evidence, covering [period]

— valid for that period only

- Paired-test results: point-in-time, conducted [date] — valid for that moment; does not establish behavior at other times
- Monitoring records: period evidence, covering [period] Completeness and known gaps:
- Available: design evidence for represented mechanism; implementation evidence for gateway configuration, identity configuration, metadata schema; operating evidence for authorization decisions and paired tests
- Known gap: no operating evidence for alternate retrieval paths (for example, direct document access via administrative interface) during the assessment period — gap documented; relates to hypothesized failure condition "alternate retrieval path that does not invoke the represented authorization mechanism"
- Known gap: no operating evidence for behavior when document authorization metadata is missing during the assessment period — gap documented; relates to hypothesized failure condition "stale, missing, or inconsistent document authorization attributes"
- Inaccessible evidence: provider-internal logs for retrieval API gateway (evidence-source state: provider-only-inaccessible) — relates to represented mechanism's internal operation; organization relies on provider-provided documentation and customer-observed authorization decision logs Evidence preservation — fictional example choices:
- Architecture record: preservation method [organization-selected method], with version/date and any integrity reference recorded
- Configuration snapshot: preserved as [structured export/snapshot method] with collection time and transformations recorded
- Retrieval authorization logs: preserved as [organization-selected log export or retention method] for the stated period, with filtering or aggregation documented
- Paired-test results: preserved as [test record/report format] with test conditions, selection basis, date, and producer recorded
- Provider-provided documentation: preserved as received with source, version/date, and any provider-supplied authenticity or provenance information recorded These are illustrative preservation choices for the fictional example, not an ODA3-required evidence package.

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence to preserve: Evidence Basis Record version and date; evidence identification by category (design, implementation, operating);

evidence source state and implementation-responsibility state for each item; provenance (producer, date, method, chain of custody, known limitations); temporal scope (point-in-time, period, superseded, reference evidence); coverage description; known gaps with reason and related assessment-question dimension; inaccessible evidence with reason and related dimension; preservation method for each item; prior Evidence Basis Record versions; references to the frozen question, boundary, applicability record, and control-understanding record.

Conclusion boundary: This Evidence Basis Record documents the evidence identified, collected, preserved, and classified for the fictional retrieval-authorization element. It records provenance, temporal scope, coverage, missing evidence, and inaccessible evidence.

It does not verify propositions against the evidence, determine whether the evidence is sufficient to support a finding, establish that the represented mechanism operated as intended, or produce a finding state.

Part VII evaluates what propositions this evidence supports and the limitations of verification; Part X forms findings from those verified results within the defined conclusion boundary.

51. Practitioner Field — Evidence Basis Record Template (Blank)

Use one Evidence Basis Record for each candidate system element whose represented control model was documented in Part V, or per assessment question where multiple elements share the same evidence basis and that basis is documented.

Assessment Question Reference: _____ System Boundary Reference: _____ Applicability Record Reference:

_____ **Control Understanding Record Reference:** _____ Candidate System Element (from Part IV/V): _____ Evidence Basis Record Version / Date: _____ Prior Evidence Basis Record Version / Reference: _____

Evidence identification — design evidence: - Item description:

_____

- **Source (producer):** _____
- **Evidence-source state:**

Customer-observed / Provider-provided / Shared-interface / Provider-only-inaccessible / Unknown: _____

- **Implementation-responsibility state:** Enterprise-controlled / Provider-controlled / Shared / Visibility-limited: _____
- Related represented-model dimension (intended outcome / mechanism / operating boundary / dependency / hypothesized failure condition):

_____ **Evidence identification — implementation evidence:**

- **Item description:** _____
- **Source (producer):** _____
- **Evidence-source state:** _____
- **Implementation-responsibility state:** _____
- **Related represented-model dimension:** _____ Evidence identification — operating evidence:
- **Item description:**

_____

- **Source (producer):** _____
- **Evidence-source state:**

_____

- **Implementation-responsibility state:** _____
- **Related represented-model dimension:** _____ Evidence provenance (for each item):
- **Producer (role, team, provider function, system):** _____
- **Date/time produced:**

_____

- Method of production (automated log, manual export, interview, observation, test execution): _____

- **Chain of custody: _____**
- **Known limitations on reliability: _____** Temporal scope and relationship (for each item):
- Point-in-time (with timestamp) / Period (with start and end) / Superseded (with successor reference) / Reference evidence (with version/date/status):

_____

- **Relationship to assessment period: _____** Coverage description:
- **Evidence identified for each represented-model dimension: _____**
- **Evidence partially represented in current basis: _____**
- **Evidence not identified or inaccessible: _____** Known gaps:
- **What evidence is missing: _____**
- **Why it is missing: _____**
- **Related assessment-question or represented-model dimension: _____**
- **Whether gap can be closed by later evidence collection: _____** Inaccessible evidence:
- **What the evidence is (to the extent describable): _____**
- **Why it is inaccessible: _____**
- **Evidence-source state: _____**
- **Related assessment-question or represented-model dimension: _____**
- **Whether alternative evidence exists: _____** Evidence preservation (for each item):
- **Preservation method used:**

_____

- **When preserved: _____**
- **Transformations applied:**

_____

- **Integrity, custody, or preservation limitations relevant to later examination: _____** Evidence-basis change provenance:
- **Material change discovered during examination: _____**
- **Source/reason for change:**

_____ - Dimensions or assessment questions affected:

_____ - Effect on verification work already performed:

_____

- Question/boundary/applicability/control-understanding revision required under §15 / §21.1 / §29.1 / §40.1: _____
- **Superseded or unreliable evidence preserved (reference): _____** Resource-planning prompts (completed by the practitioner):
- **Internal roles required for evidence identification and collection: _____**
- External specialist support required (for example, provider liaison, legal for inaccessible evidence):

_____

- **Estimated practitioner effort: _____**
- Evidence-access dependencies (for example, log-store access, provider cooperation, legal review): _____
- **Technical environment/access required: _____**
- **Scheduling constraints: _____**
- **Assumptions and uncertainties about evidence availability: _____** These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration or effort benchmarks.

Part VI Checkpoint

Can a third party who was not involved in the evidence-basis work read the Evidence Basis Record and know exactly:

- what evidence was identified, collected, or preserved; what kind of evidence each item is (design, implementation, operating);
- where each item came from and how it was preserved; what temporal scope or reference state each item represents;
- what has been identified, what is missing, and what is inaccessible; what known gaps exist and what they relate to;
- what has not yet been verified against the evidence;

— without the record having determined whether the evidence is sufficient, whether the represented mechanism operated as intended, or what finding state applies?

If not, the evidence basis is not yet sufficiently developed for later verification work.

The checkpoint is an orientation aid, not a readiness gate, assessment state or certification prerequisite.

PART VII — VERIFY CONTROL OPERATION

52. Why Verification Comes After the Evidence Basis

Part II defined the assessment question. Part III established the system boundary. Part IV identified potentially applicable criteria and candidate system elements. Part V documented the represented intended control model. Part VI built and preserved a versioned Evidence Basis Record without deciding what propositions that evidence ultimately supports.

Part VII now performs **verification**: it evaluates defined propositions against the evidence actually available in a stated Evidence Basis Record and documents what the verification supports, what contradicts it, what alternative explanations remain, and what limitations bound the result.

Verification must not silently change its own evidentiary foundation. If verification exposes the need for an additional log extract, configuration snapshot, provider record, observation, test, or other evidence item, that item must first enter a revised Evidence Basis Record under §48.1. The verification record then identifies the revised evidence-basis version used. This preserves the difference between **building the evidence basis** and evaluating propositions against it.

The clean methodological boundary for this Part is:

Part VII evaluates defined propositions against a versioned Evidence Basis Record and documents the verification procedure, evidence used, observations or analytical results, alternative explanations considered, proposition-level support, and verification limitations.

It does not convert those results into a final finding state, severity, conformity determination, certification decision, or generalized claim beyond the verified scope, period, system state, evidence basis, and procedure. Where additional evidence is required, that evidence must enter through a revised Evidence Basis Record before being used. Part X forms findings from the verified results within the defined conclusion boundary.

Verification is therefore not a second evidence-collection stage and not an early finding engine. It is the disciplined bridge between preserved evidence and later finding formation.

Part VII evaluates defined propositions against a versioned Evidence Basis Record. It records the procedure performed, evidence actually used, observations or analytical results, contradictions, corroborating relationships, plausible alternative explanations, proposition-level support and verification limitations. Evidence classification, producer, source relationship or observation method does not by itself establish the proposition. Additional evidence must first enter through a revised Evidence Basis Record. Part VII does not convert proposition-level results into a finding state, severity, conformity determination, certification decision or generalized assurance claim. Part X forms bounded findings from the verified results and recorded limitations.

53. What "Verification" Means in This Guide

A verification activity begins with a **proposition**: a bounded statement that can be evaluated against identified evidence under stated conditions.

A useful verification proposition should identify, as applicable:

26. Subject: the mechanism, configuration, event, behavior, dependency, or relationship being examined.
27. **Claimed state or behavior:** what is represented to be true.
28. Scope: the system element, path, user/authority context, environment, or other boundary condition to which the proposition applies.
29. Period or state: the relevant assessment period, point in time, version, or configuration state.
30. Evidence-basis reference: the versioned Evidence Basis Record and specific evidence items used.
31. Verification procedure: how the proposition will be evaluated using those items.
32. Limitations: known conditions that constrain interpretation of the result.

For example, a proposition may ask whether a configuration snapshot supports the represented presence of an authorization mechanism at a stated point in time, or whether operating records support that an authorization decision occurred before content retrieval for specified observed transactions during a defined period.

The proposition must not be broader than the evidence and procedure can examine.

53.1 Verification result language

Part VII records **proposition-level verification results**, not Part X finding states. Avoid using bare labels such as `Supported`, `Partial`, `Gap`, or `Unknown` as Part VII result categories because those terms could be confused with finding formation.

Prefer bounded language such as:

Verification result: The specified evidence supports proposition P-01 for the stated scope, period, configuration, and observed paths, subject to the recorded limitations.

or: Verification result: The specified evidence does not resolve proposition P-02 because the available observation path does not cover the alternate execution path identified in the Control Understanding Record.

or: Verification result: Evidence items E-04 and E-07 conflict on the configured state. The proposition remains unresolved at this verification stage pending treatment of the contradiction under the applicable later methodology.

The result should distinguish at least four things:

- what was evaluated; what evidence was actually used; what the verification procedure observed or established about the proposition;
- what remains outside the result.

Proposition-level support ≠ final finding state. Part X forms findings from the verified results and their limitations within the defined conclusion boundary.

53.2 Verification procedure is not a universal test protocol

COM-020 does not prescribe a proprietary ODA3 test suite, universal sample size, attack library, confidence threshold, or pass/fail rule.

The verification procedure should be appropriate to the proposition, system, evidence type, consequence context, and governing criterion.

Where a governing requirement, organizational methodology, sector standard, or approved test method specifies a procedure or threshold, record that source and use it within its stated scope. Do not convert an organization-specific or source-specific threshold into a universal COM-020 rule.

54. Verify Propositions About Represented Design

Design evidence may bear on propositions about represented intent, reference architecture, or applicable criterion interpretation. Verification determines what the specified design/reference evidence supports about the represented intended control model documented in Part V — intended outcome, represented mechanism, ownership/responsibility, intended operating boundary, dependencies, and hypothesized failure conditions. Evidence classification does not by itself establish that the represented design is correct, complete, or authoritative.

Design verification may examine questions such as:

- What does the specified design/reference evidence support about the relationship between the represented design recorded in Part V and the identified criterion?
- Are the represented mechanism and dependencies described consistently across authoritative design sources?
- Does the architecture representation place the mechanism on the path relevant to the assessment question?

- Do provider and enterprise design representations describe the same responsibility boundary, or do they conflict?
- Does the represented design account for the authority context, alternate path, degraded state, or dependency identified in the frozen boundary and Control Understanding Record?

A design verification result is bounded to represented intent. Even a consistent set of design records does not establish deployed configuration or operating behavior.

54.1 Source conflicts in represented design

Where design sources conflict, preserve the conflict rather than selecting the preferred representation without basis. Record:

- the conflicting evidence-item references; the versions/dates involved; which represented dimension is affected;
- whether one source supersedes another and the basis for that conclusion; whether the conflict affects later implementation or operating verification.

If resolving the conflict requires new evidence, update the Evidence Basis Record under §48.1 before using that evidence.

55. Verify Propositions About Implemented State

Implementation evidence may bear on propositions about configured or deployed state. Verification determines what the specified implementation evidence supports about that state, for the stated version, environment, time and evidence basis, and its relationship to the represented design. Evidence classification does not by itself establish that the recorded configuration was complete, representative, or actually deployed during the assessment period.

Relevant propositions may concern:

- whether the identified mechanism is present in the examined configuration; whether the configuration corresponds to the version, environment, and system element within scope;
- whether identity, authority, tool, retrieval, policy, monitoring, or other relevant settings match the represented configuration state;
- whether required dependencies are connected or configured as represented; whether an alternate or bypass path exists in the examined configuration;
- whether a configuration change occurred during the assessment period and which state the evidence represents.

Implementation verification must preserve temporal boundaries. A snapshot at one time supports only propositions that the snapshot and its provenance can bear on. It does not automatically establish the configuration for the entire assessment period.

Likewise, the presence of a configured mechanism does not establish that it executed on every relevant path, operated correctly, resisted bypass, or produced the intended outcome. Those are operating propositions.

55.1 Relationship between represented design and implemented state

A verification record should state the relationship actually examined rather than collapsing design and implementation into one conclusion.

Examples include:

- configuration evidence corresponds to the represented mechanism for the stated version and environment; configuration evidence differs from the represented design in

identified respects;

- available evidence does not establish whether the deployed state corresponds to the represented design; multiple deployed states existed during the assessment period and require separate proposition treatment.

None of these statements is itself a Part X finding.

56. Verify Propositions About Operation

Operating evidence may bear on propositions about events or behavior during a defined period or condition. Verification determines what the specified operating evidence supports within that observation boundary; it does not treat the evidence classification itself as proof that the recorded behavior is complete or representative. Operating propositions should be narrow enough to preserve the difference between observed operation and generalized effectiveness.

Operating propositions should be narrow enough to preserve the difference between **observed operation** and generalized effectiveness.

Examples include whether, for specified observed transactions or controlled tests:

- an authorization decision occurred before a consequential action; an unauthorized request was refused or constrained;
- a tool invocation corresponded to the authority represented at execution time; a monitoring or escalation event was generated when the represented trigger condition occurred;
- an expected record was produced for an observed action; an alternate path bypassed the represented mechanism; a dependency failure changed the observed behavior;
- a stale identity, policy, configuration, or authority state affected the execution path.

The verification record must state the observation path and conditions.

Evidence from one path, user class, configuration, model version, provider state, or test condition does not automatically establish behavior for another.

56.1 Observation and testing

Where direct observation or testing is part of the Evidence Basis Record, verification should document:

- the proposition being evaluated; test or observation reference and version; environment and relevant system state;
- identities, authority context, inputs, or preconditions relevant to interpretation; selection basis for cases or paths examined;
- expected represented behavior, where applicable; observed result, including refusals, failures, unsuccessful attempts, and unexpected behavior;
- known interference, instrumentation, or observation limitations; repeatability information where relevant; relationship to the assessment period and production state.

COM-020 does not prescribe a universal number of tests or observations.

Selection should be reasoned and documented in relation to the proposition and consequence boundary.

57. Corroboration, Contradiction, and Alternative Explanations

Distinct items, producers, or observation paths must have their relationships explicitly recorded. Corroboration must not imply independent assurance.

Verification should not stop when one evidence item appears consistent with the proposition. It should consider whether other evidence corroborates, narrows, contradicts, or provides an alternative explanation for the observed result.

Corroboration means that distinct evidence items or observation paths bear consistently on the same proposition. Corroboration does not require that every item have the same producer or source relationship, and it does not create an automatic hierarchy in which one source class always outranks another.

For each proposition requiring corroboration under the examination plan, consider:

- whether design, implementation, and operating evidence tell a consistent story; whether the same result is visible through more than one relevant observation path;
- whether provider-provided and customer-observed evidence align or conflict; whether temporal differences explain apparent inconsistencies;

- whether the observed result could be caused by a different mechanism than the one represented; whether a bypass, stale state, fallback, cache, alternate route, hidden dependency, or human intervention could explain the result;
- whether instrumentation itself could have changed or obscured behavior.

57.1 Testing alternative explanations

Part V identified hypothesized failure conditions. Part VII uses those hypotheses as prompts for verification, not as presumptions that failure occurred.

For a proposition that appears supported, ask:

What plausible alternative explanation could produce the same observed result without the represented mechanism having operated as claimed?

For a proposition that appears contradicted, ask:

Could the apparent contradiction arise from temporal mismatch, configuration drift, incomplete observation, different authority context, alternate path, evidence transformation, or another bounded explanation?

Where an alternative explanation can be evaluated using the current Evidence Basis Record, document the procedure and result. Where it requires new evidence, route that evidence through §48.1 before using it.

The objective is not to eliminate every imaginable alternative explanation. It is to identify and test plausible alternatives that bear on the proposition within the defined assessment question and consequence boundary.

58. Refusals, Negative Evidence, and Verification Limitations

Part VII records how these conditions affect verification of the specific proposition. Part IX determines how negative and inconclusive verification results are preserved and treated across the broader assessment record.

Part VII may encounter evidence that is absent, contradictory, inaccessible, refused, or negative. This section records how those conditions affect a **specific verification proposition**. It does not replace Part IX, which owns the broader methodology for negative and inconclusive results.

Keep the following states distinct:

- Observed negative evidence: an observed event or state inconsistent with the proposition or represented behavior.
- Checked-and-not-observed: a defined observation or search was performed and the expected event/state was not observed within the stated conditions.
- Evidence not identified: the Evidence Basis Record contains no identified item for the relevant proposition or path.
- Evidence inaccessible: an identified item exists or is represented to exist but cannot be accessed.
- Evidence refused: a relevant party declines to provide an identified or requested item; record the refusal, scope, authority/context, and any stated reason.
- Unknown: the relevant state cannot be determined from the available basis and procedure.
- Contradictory evidence: two or more evidence items bear inconsistently on the same proposition and the contradiction is not resolved within the verification performed.

These conditions are not interchangeable. In particular, absence of evidence ≠ evidence of absence**, and** Unknown ≠ Checked-and-not-observed.

A refusal does not automatically prove that the underlying control failed. Inaccessible provider evidence does not automatically establish a Gap. Conversely, provider representations do not automatically resolve a proposition merely because direct evidence is inaccessible. Record the verification limitation precisely and carry the result forward.

#Checked-and-not-observed requires a defined observation or search method, scope, conditions, and result; the absence of an item from the Evidence Basis Record alone does not establish checked-and-not-observed.

58.1 When Verification Changes During Examination

Verification may change when the evidence basis changes, a procedure is corrected, a contradiction is resolved, a new system state is discovered, or an earlier result is found to have relied on an invalid assumption.

When a Verification Record changes materially:

- preserve the prior Verification Record rather than overwriting it; issue a revised version identifying what changed, why, and the source or trigger for the change;
- identify the proposition(s) affected; identify the Evidence Basis Record version and evidence items used before and after the change;
- identify any changed procedure, environment, system state, assumption, observation path, or analytical method;
- state whether the earlier verification result remains usable within its original boundary, is superseded for later work, or requires separate treatment;
- determine whether the assessment question, system boundary, applicability record, Control Understanding Record, or Evidence Basis Record requires revision under §15, §21.1, §29.1, §40.1, or §48.1; and preserve contradictions, unsuccessful attempts, negative

observations, and superseded results rather than silently deleting them.

A later Verification Record does not retroactively change the evidence basis, procedure, system state, or observation conditions against which earlier verification was performed. Later finding formation must identify which verification version it relies on.

59. Worked Example — Verifying Control Operation (Illustrative)

Note: *Illustrative fictional system — system names, record identifiers, configurations, evidence items, procedures, and results below are invented solely to demonstrate the method. They are not ODA3 benchmarks, prescribed test protocols, representations of any real system/provider, or ODA3 assessment outcomes.*

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence Basis Record used: [fictional Evidence Basis Record version/date]. No evidence outside that version is used unless first incorporated through a revised Evidence Basis Record under §48.1.

Proposition P-01 — represented design

Proposition: The represented design places a retrieval-authorization decision on the retrieval path before content is returned downstream.

Evidence used: fictional architecture record [ARCH-RET-01]; fictional internal policy [SEC-POL-014 §4.2]; fictional provider documentation [reference].

Procedure: compare the path and authorization sequence represented in the cited design/reference evidence; identify version/date conflicts and responsibility-boundary differences.

Verification result: The cited fictional design records consistently represent an authorization decision before downstream content return for the documented retrieval path. This result supports only the represented-design proposition. It does not establish deployed configuration or actual operation.

Limitations / alternative explanations: [fictional unresolved design conflict or none identified within the stated records];

alternate retrieval paths are treated separately.

Proposition P-02 — implemented state

Proposition: The examined configuration snapshot contains the represented authorization-enforcement setting for the stated environment and configuration state.

Evidence used: fictional configuration export [reference]; fictional identity role-assertion configuration [reference]; fictional metadata schema [reference].

Procedure: inspect the preserved configuration items and map them to the represented mechanism and relevant dependency references.

Verification result: The fictional snapshot contains configuration elements corresponding to the represented authorization mechanism for the stated point in time. This does not establish that the same configuration persisted for the full assessment period or that it executed on every relevant path.

Limitations / alternative explanations: configuration state outside the snapshot time remains bounded by available change/history evidence;

provider-internal configuration is not directly observed.

Proposition P-03 — observed operation on tested paths

Proposition: For the practitioner-defined document/user combinations included in the fictional controlled exercise, the observed retrieval path refused content where the requesting identity was represented as lacking direct-access authorization.

Evidence used: fictional paired-test record [reference]; fictional authorization-decision logs [reference]; fictional monitoring records [reference].

Procedure: compare each documented test identity/document condition with the corresponding retrieval result and available authorization-decision record; preserve unsuccessful attempts and anomalous results.

Verification result: For the fictional tested combinations recorded in the stated exercise, the observed results are consistent with the proposition for those tested paths and conditions, subject to the limitations below.

Limitations / alternative explanations: the exercise does not establish behavior for untested identities, documents, alternate retrieval paths, missing authorization metadata, other configurations, or other periods. Provider-internal processing is not directly observed.

A consistent observed refusal does not by itself prove which internal mechanism caused the refusal unless the available evidence supports that causal attribution.

Proposition P-04 — alternate-path hypothesis

Proposition: Available evidence resolves whether an alternate retrieval path can return content without invoking the represented authorization mechanism.

Evidence used: [fictional evidence-item references].

Procedure: examine the Evidence Basis Record for architecture, configuration, logs, or test evidence bearing on alternate retrieval paths.

Verification result: The current fictional Evidence Basis Record does not resolve the proposition for the identified alternate path.

Additional evidence, if obtained, must first be incorporated into a revised Evidence Basis Record before a later verification version uses it.

Limitations: unresolved alternate-path coverage; no inference is made that bypass either exists or does not exist.

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence to preserve: Verification Record version/date; proposition identifiers and wording; Evidence Basis Record version;

evidence-item references actually used; verification procedure and version; system/environment/configuration state; observation or test conditions; selection basis; observed results including refusals, failures, unsuccessful attempts, anomalies and contradictions;

alternative explanations considered and how they were treated; verification limitations; prior Verification Record versions;

references to the frozen question, boundary, applicability record, Control Understanding Record, and Evidence Basis Record.

Conclusion boundary: These fictional verification results establish only what the specified evidence and procedures support about the defined propositions for the stated system state, period, paths, and conditions. They do not establish generalized control effectiveness beyond those boundaries, legal or regulatory conformity, certification, or a final finding state. Unresolved alternate paths and inaccessible provider-internal evidence remain explicit limitations. Part X forms any finding from the verified results within the defined conclusion boundary.

60. Practitioner Field — Verification Record Template (Blank)

Use one Verification Record per proposition or a grouped record where multiple propositions genuinely share the same evidence basis, procedure, scope, and conclusion boundary and the grouping is explicit.

Assessment Question Reference: _____\ System Boundary Reference: _____\ Applicability Record Reference: _____\ Control Understanding Record Reference: _____\ Evidence Basis Record Version / Date: _____\ Verification Record Version / Date: _____\ Prior Verification Record Version / Reference: _____

Proposition

- Proposition ID: _____
- Proposition statement: _____
- Subject / mechanism / behavior: _____
- Scope / path / authority context: _____
- Relevant period / point in time / configuration state: _____
- Criterion or represented-model dimension to which proposition relates: _____

Evidence actually used

- Evidence item reference(s): _____
- Evidence classification(s): Design / Implementation / Operating:

_____

- Evidence-source state(s): _____
- Producer/source relationship(s): _____
- Temporal relationship to proposition: _____
- Relevant provenance/integrity/access limitations: _____

Verification procedure

- Procedure or analytical method: _____
- Procedure source/version, if applicable: _____
- Environment / system state: _____
- Observation path / instrumentation: _____
- Selection basis for cases/paths examined: _____
- Inputs / identities / authority context / preconditions:

_____

- Expected represented behavior, where applicable: _____
- Deviations from planned procedure: _____

Observations and analytical results

- Observed result(s): _____
- Refusals / negative observations: _____
- Unsuccessful attempts / test failures: _____
- Anomalies: _____
- Contradictory evidence: _____
- Checked-and-not-observed state(s), with conditions: _____
- Unknown / inaccessible / refused evidence relevant to proposition:

_____

Corroboration and alternative explanations

- Corroborating evidence or observation paths: _____
- Conflicting evidence or observations: _____
- Plausible alternative explanations considered: _____
- Alternative explanations evaluated using current Evidence Basis Record: _____
- Alternative explanations requiring additional evidence: _____
- Evidence Basis Record revision reference, if additional evidence was added under §48.1: _____

Proposition-level verification result

- Bounded verification result in prose: _____
- Scope/period/configuration/path to which result applies:

_____

- Verification limitations: _____
- What the result does **not** establish: _____
- Unresolved propositions or conditions carried forward: _____

Do not use this field to assign a Part X finding state, severity grade, conformity determination, certification decision, or mandated remediation timeline.

Verification-change provenance

- Material verification change discovered: _____
- Source/reason for change: _____
- Proposition(s) affected: _____
- Prior Evidence Basis Record version: _____
- Current Evidence Basis Record version: _____
- Procedure/system-state/assumption change: _____
- Effect on prior verification result: _____
- Earlier result preserved at reference: _____
- Upstream revision required under §15 / §21.1 / §29.1 / §40.1 / §48.1: _____

Resource-planning prompts (completed by the practitioner)

- Internal roles required for verification: _____
- External specialist support required: _____
- Estimated practitioner effort: _____
- Environment/test access dependencies: _____
- Provider or third-party cooperation dependencies: _____
- Legal/privacy/safety review dependencies, where applicable:

_____

- Scheduling constraints: _____
- Assumptions and uncertainties: _____

These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration, sample-size, testing, or effort benchmarks.

Part VII Checkpoint

Can a third party who was not involved in the verification read the Verification Record and determine:

- the exact proposition evaluated; the Evidence Basis Record version and evidence items actually used; the procedure and system/observation conditions;
- the observations, negative results, contradictions, and unsuccessful attempts preserved; the alternative explanations considered;
- the bounded proposition-level result; the limitations and unresolved conditions; whether additional evidence entered through a revised Evidence Basis

Record rather than being silently added;

- and what the verification explicitly does not establish;

— without the record having converted the result into a final finding, severity, conformity decision, certification decision, or generalized claim?

If not, the verification record is not yet explicit in its proposition, evidence basis, procedure, result, limitations, and conclusion boundary for later finding formation.

The checkpoint is an orientation aid, not a readiness gate, assessment state, certification prerequisite, or ODA3 decision rule.

PART VIII — ASSESS AI-SPECIFIC BEHAVIOUR

61. Why AI-Specific Behaviour Requires Separate Examination

AI systems may exhibit probabilistic and model-mediated behavior, dependency on training data and retrieval context, delegated authority to invoke tools or influence other systems, and sensitivity to model, provider, configuration, or input-distribution changes. These properties create examination considerations distinct from purely deterministic logic, though non-AI systems can also exhibit nondeterminism due to distribution, concurrency, or adaptation.

Part VIII specializes the Part VII proposition-verification method for AI-specific behavioural dimensions. It does not create a parallel verification layer or result taxonomy.

The clean methodological boundary for this Part is:

Part VIII applies the proposition-level verification discipline established in Part VII to AI-specific behavioural dimensions relevant to the frozen assessment question, system boundary, applicability determination and represented control model. Examination is performed against a versioned Evidence Basis Record. Where additional evidence is required, it first enters a revised Evidence Basis Record under §48.1. Part VIII records what the specified evidence and procedure support about the AI-specific proposition and the limitations of that examination. It does not assign universal thresholds, materiality, acceptability, finding states, severity, conformity, certification status, functional-safety status or generalized assurance. Part IX governs the broader treatment of negative and inconclusive results; Part X forms bounded findings from verified results and recorded limitations.

62. Probabilistic Behaviour and Output Variability

Probabilistic or model-mediated behaviour means that, for varied valid inputs or repeated trials, the system may produce differing outputs, including differing adherence to constraints, differing factual assertions, or differing levels of detail.

Examination may consider bounded propositions such as:

- **Instruction following:** What do specified observations support about adherence to stated constraints (e.g., output format, length, negative constraints) under stated input conditions?
- **Unsupported generation / contradiction / variability:** What do specified observations support about generation that is unsupported by provided context, contradicts provided context or ground truth, varies across semantically similar inputs, or exhibits calibration properties?
- **Calibration and consistency:** What do specified observations support about the relationship between expressed confidence or probabilistic output and observed accuracy or consistency, where such properties are represented?

These dimensions — unsupported generation, contradiction, variability and calibration — are separate potentially applicable dimensions. COM-020 does not prescribe a universal definition, metric, or threshold for any of them. Where a governing criterion or organizational methodology defines a threshold or measurement method, record its source, version, and scope.

Boundary discipline: COM-020 does not prescribe universal hallucination rates, accuracy percentages, or pass/fail thresholds.

Proposition results are bounded to tested conditions, observation paths, and Evidence Basis Record version.

63. Retrieval-Augmented Generation (RAG) System Behaviour

RAG systems introduce relationships between retrieval behavior, generation grounded in retrieved context, citation/source correspondence, and authorization/trust-boundary behavior. These are distinct dimensions and are not interchangeable failure modes.

Examination may consider bounded propositions such as:

- **Retrieval behavior:** What do specified observations support about which documents or passages were retrieved for stated queries, ranking/cutoff behavior (including applicable top-k or similarity-threshold configuration), and coverage of relevant sources under stated retrieval configuration?
- **Grounding / claim-source relationship:** What do specified observations support about whether generated claims are supported by retrieved context versus extrinsic or unsupported claims?
- **Citation / source correspondence:** What do specified observations support about whether provided citations or source references accurately correspond to the specific retrieved material supporting the claim?
- **Authorization / trust-boundary behavior:** What do specified observations support about retrieval or exposure of content relative to the requesting identity's represented authorization and the applicable identity/authorization boundary? Retrieval-configuration limits are recorded separately under retrieval behavior and its limitations.

Retrieval failure, grounding failure, citation failure, and authorization-boundary behavior are distinct. Part VIII preserves their separation; Part VII evaluates what specified evidence supports for each.

Boundary discipline: Groundedness and citation correspondence are examined as behavioural properties against the Evidence Basis Record, not as universal scoring metrics.

64. Agentic Systems and Tool Use

Agentic systems may be represented as having delegated capability to invoke external tools, APIs, or actions based on internal reasoning. Examination focuses on represented authority, delegated capability, execution boundaries, current authority at action time, intervention/escalation and observable tool behavior.

Examination may consider bounded propositions such as:

- **Represented authority and delegated capability:** What authority and tool capabilities are represented as delegated to the agent for the stated version and environment?
- **Execution boundaries and current authority at action time:** What do specified observations (e.g., execution logs, authorization checks, policy enforcement records) support about whether tool invocations occurred within represented execution boundaries and with current authority at the time of action?
- **Parameter and context handling:** What do specified observations support about validation or handling of parameters and authority context passed to external tools?
- **Intervention, escalation and human oversight interface:** What do specified observations support about intervention, escalation, or human-oversight interface behavior as represented for agentic actions?

Whether authority is "excessive" requires an identified criterion. COM-020 does not label authority as excessive without such a criterion.

Boundary discipline: Presence of an allowlist or delegation description does not by itself establish that tool-use behavior occurred within bounds. Verification determines what specified evidence supports within the defined observation period, paths, and Evidence Basis Record version.

65. Model, Provider and Relevant Behavioural Change

AI systems may be subject to changes in input distribution, model/version, provider service, configuration, dependency, or the relationship between inputs and expected outputs. Every observed change is not automatically "drift."

Examination requires applicability and proposition definition. Distinguish, where relevant:

- **Input / distribution change:** change in input data or distribution relative to prior state;
- **Relationship / performance change:** change in the relationship between inputs and outputs or in observed performance under defined conditions;
- **Model / version change:** change in model version, weights, prompt, retrieval index, or associated artifact;
- **Provider-service change:** change in third-party model API, platform, or service behavior;

- **Configuration / dependency change:** change in configuration, dependency, or environment.

Examination may consider bounded propositions such as:

- **Change-detection mechanisms:** What do specified observations support about monitoring that may bear on detection of relevant changes?
- **Provider-service change impact:** What do specified observations support about impact of provider-service changes on represented behavior, where such changes occurred?
- **Represented re-examination / reverification triggers:** Where governing criteria or the represented control model define conditions under which re-examination or reverification occurs, what do specified observations support about whether those triggers were represented and whether they were observed?

Use **represented re-examination/reverification triggers** where such triggers exist in governing criteria or represented control model. COM-020 does not create an undefined lifecycle requirement.

Boundary discipline: Part VIII does not mandate specific drift-detection algorithms or timelines. It examines what evidence supports about represented mechanisms and observed behavior during the assessment period and Evidence Basis Record version.

66. Physical Consequence Paths

Where an AI system can initiate, modify, authorize, inhibit, or otherwise influence physical action (e.g., autonomous vehicles, robotics, industrial control systems, medical devices), the examination must explicitly address the physical consequence path.

Examination may consider bounded propositions such as:

- **Path identification:** What is the sequence from digital inference or planning to physical command and actuator?
- **Mechanisms represented as constraining, interrupting, authorizing or placing the physical action into a defined state:** What mechanisms are represented as constraining, interrupting, requiring authorization for, or placing physical action into a defined safe or bounded state?
- **Authority and oversight for physical action:** What authority is required before physical action is executed, and how is that authority enforced and observed?

Boundary discipline: COM-020 includes examination of the physical consequence path and evidence relevant to assessment question. It does **not** make a functional-safety determination, grant sector-specific safety approval, or authorize physical return-to-service. Functional-safety determination and return-to-service authority remain with applicable safety regime and authorized decision authority.

Where relevant ODA3 control context from Physical AI Security Framework (PAI-SF™) is applicable because assessment basis explicitly identifies it, it may be referenced with provenance. COM-020 does not mandate PAI-SF™ as the sole method unless assessment basis explicitly makes it applicable.

67. Human Oversight Models

Human oversight may be an applicable examination dimension where it forms part of the represented control model or identified criterion. Examination focuses on represented oversight arrangements — human-in-the-loop, human-on-the-loop, human-in-command — as represented arrangements, not as universal assurance categories.

Examination may consider bounded propositions such as:

- **Human-in-the-loop (HITL):** What do specified observations support about whether explicit human approval was represented as required before a consequential action and whether that approval was recorded within stated observation conditions?
- **Human-on-the-loop (HOTL):** What do specified observations support about whether human monitoring and intervention or override capability was represented and observable within stated conditions?
- **Human-in-command (HIC):** What do specified observations support about whether a human retains represented authority, decision rights, intervention rights, and accountability assignments as

defined in the represented control model over system objectives and deployment context as represented?

Boundary discipline: Examination verifies what specified evidence supports about oversight mechanism as represented (e.g., approval logs, override interface, accountability record). It does not evaluate subjective quality of human decision-making beyond what specified evidence supports.

Part VIII may record negative, contradictory, checked-and-not-observed, unresolved, or inconclusive conditions arising from AI-specific proposition-level examination. It records how those conditions affect the specific proposition and preserves the applicable evidence, procedure, scope, observation conditions, and limitations. **Part IX owns the broader methodology for preserving and treating negative and inconclusive results across the assessment record.** Part VIII does not convert those conditions into finding states.

68. Worked Example — Assessing AI-Specific Behaviour (Illustrative)

Note: *Illustrative fictional system — system names, record identifiers, configurations, evidence items, procedures, and results below are invented solely to demonstrate method. They are not ODA3 benchmarks, prescribed test protocols, representations of any real system/provider, or ODA3 assessment outcomes.*

Assessment question (from Part II): For Support Assistant A in defined production configuration and operating period, what evidence supports claim that retrieval-time authorization prevents user from receiving content they would not be permitted to access directly, and that system's RAG behaviour remains grounded in retrieved documents?

Evidence Basis Record used: [fictional Evidence Basis Record version/date]. **Verification Record used:** [fictional Verification Record version/date]. No evidence outside those versions is used unless first incorporated through revised Evidence Basis Record under §48.1.

AI-Specific Behaviour Examination Record — fictional represented behaviour:

Probabilistic behaviour and output variability:

- Proposition examined: For stated out-of-scope prompt set executed under stated conditions, what do observations support about refusal behavior?
- Evidence used: Fictional red-team prompt suite (version reference) executed against production endpoint during defined period; Evidence Basis Record reference.
- Observation: For fictional tested prompt set under stated conditions, recorded results show refusal in cases represented as out-of-scope, with a subset of cases showing evasive but non-committal responses that did not disclose unauthorized information within observation scope.
- Limitation: Prompt suite does not cover all possible linguistic variations; observation bounded to tested conditions, paths, and Evidence Basis Record version.

RAG behaviour:

- Retrieval behavior: For practitioner-defined document/user combinations included in fictional controlled exercise, what do observations support about documents retrieved?
- Evidence used: Fictional paired-test record [reference]; retrieval logs [reference].
- Observation: For tested combinations and stated retrieval configuration, specified retrieval logs show retrieval of [specified fictional document set] matching represented ranking/cutoff behavior.
- Limitation: Provider-internal retrieval processing not directly observed; instrumentation may be incomplete.
- Grounding / claim-source relationship: For tested combinations, what do observations support about grounding of generated claims in retrieved context?
- Observation / analytical result: For the tested claim/source pairs and stated observation conditions, the specified fictional comparison procedure recorded whether each examined generated claim had corresponding support in the retrieved text within the defined comparison scope.
- Citation / source correspondence: For cases where citations were provided, what do observations support about correspondence between citation and retrieved material?
- Observation / analytical result: For the tested citation/source pairs and stated observation conditions, the specified fictional comparison procedure recorded whether each examined citation referenced the retrieved passage identified as bearing on the associated claim within the defined comparison scope.

- Authorization / trust-boundary behavior: For tested identities and documents, what do observations support about retrieval relative to authorization?
- Observation: For tested combinations, specified retrieval records show refusal or no retrieval record for cases represented as lacking direct-access authorization within stated observation paths. This does not establish absence of unobserved retrieval paths.
- Limitation: Unobserved paths, network-layer blocking, or incomplete instrumentation remain possible alternative explanations.

Agentic systems and tool use:

- Proposition examined: For stated period and tool allowlist [orders-readonly-v3], what do observations support about tool invocations?
- Evidence used: Fictional tool invocation logs for defined period; Evidence Basis Record reference.
- Observation: For recorded tool-invocation logs within the stated period and observation path, the records associate the observed invocations with [allowlist schema reference], [service identity reference], and the represented authority context recorded for those events.
- Limitation: The recorded service identity and authority context do not by themselves establish that authority was current at action time. Whether current authority is established depends on the applicable authorization evidence and verification procedure. Logs also do not capture invocations that may have been blocked at the network layer before reaching the application logging pipeline.

Model, provider and relevant behavioural change:

- Proposition examined: What do observations support about represented re-examination triggers and monitoring that may bear on behavioural change detection?
- Evidence used: Fictional monitoring dashboard configuration and represented trigger documentation; Evidence Basis Record reference.
- Observation: Configuration evidence records monitoring and represented re-examination/reverification trigger definitions as [fictional description]. Operating evidence records monitoring activity for the defined period. For [specified fictional model update reference], a re-examination/reverification record was **not identified in the stated Evidence Basis Record**.
- Limitation: Evidence not identified is not the same as checked-and-not-observed. Unless a defined search or observation method, scope, conditions, and result are recorded, this examination does not establish that re-examination/reverification did not occur. Baseline definition and provider-service changes remain bounded to the available evidence and observation paths.

Physical consequence paths:

- Proposition examined: What does the represented system boundary and specified evidence support about whether the system can initiate or influence physical action?
- Observation: Not applicable. Support Assistant A is text-based retrieval system with no physical actuation capabilities as represented.

Human oversight models:

- Proposition examined: What do observations support about represented human oversight for high-consequence actions?
- Evidence used: Fictional approval log [reference]; oversight interface description [reference].
- Observation: For [specified fictional high-consequence actions] within stated observation conditions, approval records show human approval recorded under [identity/context reference].

Assessment question: For Support Assistant A in defined production configuration and operating period, what evidence supports claim that retrieval-time authorization prevents user from receiving content they would not be permitted to access directly, and that system's RAG behaviour remains grounded in retrieved documents?

Evidence to preserve: AI-Specific Behaviour Examination Record version and date; Verification Record version and date; Evidence Basis Record version and date; propositions examined for each AI-specific dimension; evidence items used; observed behavioural results including anomalies and unsuccessful attempts; limitations of observation scope; alternative explanations for observed variability; references to frozen question, boundary, applicability, control understanding, evidence basis and verification records.

Conclusion boundary: This examination record documents proposition-level examination results concerning AI-specific behavioural dimensions of fictional Support Assistant A for stated system state, period, conditions, Evidence Basis Record and Verification Record versions. It records what the specified

evidence and procedures support regarding probabilistic output, RAG behavior (retrieval, grounding, citation, authorization), tool use, behavioural change monitoring, and human oversight. It does not establish generalized control effectiveness beyond those boundaries, legal or regulatory conformity, certification, functional safety, or final finding state. Unresolved behavioural variations and inaccessible provider-internal evidence remain explicit limitations. Part X forms any finding from these verified behavioural results within defined conclusion boundary.

69. Practitioner Field — AI-Specific Behaviour Examination Record (Blank)

Use this record to document examination of AI-specific behavioural dimensions for a candidate system element. This record specializes Part VII proposition-verification; it does not create parallel finding taxonomy.

Assessment Question Reference: _____

System Boundary Reference: _____ Control Understanding Record Reference: _____

Evidence Basis Record Version / Date: _____

Verification Record Reference / Version / Date: _____ AI-Specific Behaviour Record Version / Date: _____ Dimension 1: Probabilistic Behaviour and Output Variability

- Proposition examined: _____
- Evidence used (reference to Evidence Basis Record): _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____
- Relationship to Verification Record: _____

Dimension 2: RAG Behaviour (if applicable) — separate propositions for retrieval, grounding, citation, authorization

- Retrieval behavior — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____
- Grounding / claim-source relationship — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations: _____
- Citation / source correspondence — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations: _____
- Authorization / trust-boundary behavior — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____ Dimension 3: Agentic Systems and Tool Use (if applicable)
- Represented authority / delegated capability — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____
- Execution boundaries / current authority at action time — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations: _____
- Parameter / context handling — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Intervention / escalation / human-oversight interface — proposition examined: _____

Dimension 4: Model, Provider and Relevant Behavioural Change (if applicable)

- Change type applicability (input/distribution, relationship/performance, model/version, provider-service, configuration/dependency): _____
- Proposition examined (change-detection, provider-service impact, represented re-examination triggers): _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____ Dimension 5: Physical Consequence Paths (if applicable)
- Path identification — proposition examined: _____
- Mechanisms represented as constraining/interrupting/authorizing/placing into defined state — proposition examined: _____
- Authority and oversight for physical action — proposition examined: _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____
- *Note: COM-020 examines path and relevant evidence. Functional-safety determination and return-to-service authority remain with applicable safety regime and authorized decision authority. PAI-SF™ may be referenced only where assessment basis explicitly makes it applicable.*

Dimension 6: Human Oversight Models (if applicable)

- Proposition examined (HITL, HOTL, HIC as represented): _____
- Evidence used: _____
- Observation / Analytical result: _____
- Limitations and alternative explanations: _____ Summary of Behavioural Limitations:
- Unresolved behavioural variations or conditions carried forward: _____
- What this behavioural examination does not establish: _____ Resource-planning prompts (completed by the practitioner):
- Internal roles required for behavioural examination: _____
- External specialist support required (e.g., red team, data science): _____
- Estimated practitioner effort: _____
- Testing/environment access dependencies: _____
- Assumptions and uncertainties about behavioural observation: _____

These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration, sample-size, testing, or effort benchmarks.

Part VIII Checkpoint

Can a third party who was not involved in behavioural examination read this record and determine:

- which AI-specific behavioural dimensions were examined; specific propositions tested for each dimension;
- evidence actually used and observations made, bounded to Evidence Basis Record version and observation conditions;
- limitations, alternative explanations, and unresolved variations; that no universal thresholds, scores, or findings were generated;
- that negative, contradictory, checked-and-not-observed, unresolved, and inconclusive results are preserved with their applicable evidence, procedure, scope, conditions, and limitations for Part IX;

— without the record having converted behavioural observations into a final finding, severity grade, conformity decision, or generalized assurance claim?

If not, revise the AI-Specific Behaviour Examination Record so that its propositions, evidence basis, Verification Record relationship, procedure, result, limitations, and conclusion boundary are explicit before it is carried into later finding formation.

The checkpoint is an orientation aid, not a readiness gate, assessment state, certification prerequisite, or ODA3 decision rule.

PART IX — HANDLE NEGATIVE AND INCONCLUSIVE RESULTS

70. Why Negative and Inconclusive Results Require Separate Treatment

Verification can produce conditions that cannot be represented faithfully by a single affirmative/negative label. An observed negative event, a contradiction, a defined search in which an expected item was not observed, an unidentified evidence item, inaccessible or refused evidence, and an unresolved analytical result describe different evidentiary or verification conditions.

Part IX preserves those distinctions across the assessment record. It does not infer that a proposition is false merely because support was not established, and it does not convert an unresolved condition into a finding state.

The clean methodological boundary for this Part is:

Part IX governs the preservation and treatment of negative, contradictory, checked-and-not-observed, evidence-not-identified, inaccessible, refused, unresolved, and inconclusive conditions already arising from the versioned Evidence Basis Record and proposition-level verification in Parts VII and VIII. It records the affected proposition, evidence basis, procedure or search method where applicable, scope, conditions, result, provenance, contradictions, alternative explanations, and limitations. If additional evidence is required, it first enters a revised Evidence Basis Record under §48.1 and is used only through an applicable revised verification record. Part IX does not create a parallel evidence methodology, determine evidence sufficiency, assign materiality or severity, resolve finding states, or make conformity, certification, functional-safety, or generalized assurance conclusions. Part X forms bounded findings from the verified results and preserved limitations within the defined conclusion boundary.

71. What "Negative and Inconclusive" Means in This Guide

For this guide, preserve the following states according to what the record actually establishes:

- **Observed negative evidence:** An observed event or state inconsistent with the proposition or represented behaviour within the stated observation conditions.
- **Contradictory evidence:** Distinct evidence items or observation paths bear inconsistently on the same proposition, and the applicable verification record does not resolve that inconsistency.
- **Checked-and-not-observed:** A defined observation or search method was performed and the expected event, state, or item was not observed within the recorded scope, conditions, and result. Absence from the Evidence Basis Record alone does not establish this state.
- **Evidence not identified:** No relevant evidence item has been identified in the stated Evidence Basis Record for the proposition or path at issue.
- **Evidence inaccessible:** A relevant item is identified or represented as existing, but the stated assessment process cannot access it within the recorded boundary.
- **Evidence refused:** A relevant party explicitly declines to provide an identified or requested item; record the refusal and its scope without inferring the item's contents.
- **Unknown:** A relevant state cannot be determined from the available evidence basis and applicable procedure.
- **Unresolved / inconclusive verification result:** The applicable verification procedure produces a result that does not resolve the proposition within the stated evidence basis, procedure, scope, and conditions.

These states may describe different aspects of the same proposition or evidence path. Do not force them into a single mutually exclusive label where more than one state is needed to represent the record accurately.

72. Preserving Negative and Inconclusive Results

Part IX does not rewrite the proposition-level result recorded in Part VII or VIII. It preserves that result and the conditions needed for later interpretation.

For each applicable state, record:

33. **Affected proposition and source record:** Reference the applicable Verification Record or AI-Specific Behaviour Examination Record, including version/date.
34. **State actually established:** Record one or more applicable states from §71 without collapsing distinct conditions.
35. **Evidence basis and provenance:** Identify the Evidence Basis Record version and the evidence items or observation paths that bear on the state.
36. **Procedure, scope, and conditions:** For checked-and-not-observed or other procedure-dependent states, record the actual search/observation method, scope, period, conditions, and result.
37. **Contradictions and alternative explanations:** Preserve unresolved conflicts and plausible alternatives relevant to interpretation without selecting one unless the applicable evidence and verification procedure support doing so.
38. **Limitations and handoff:** Record what remains unresolved and carry the preserved state and limitations into Part X without assigning a finding state.

If later work identifies additional evidence, that evidence first enters a revised Evidence Basis Record under §48.1. Any changed proposition-level result must be produced through the applicable revised verification record rather than by editing the Part IX treatment record to imply that verification occurred.

73. Distinguishing Negative Evidence States

The following distinctions are mandatory to the integrity of the record:

- **Observed negative evidence ≠ generalized evidence of absence.** An observed negative event establishes what was observed under the stated conditions; it does not by itself establish absence across unobserved conditions, paths, states, or periods.
- **Checked-and-not-observed ≠ evidence not identified.** Checked-and-not-observed requires a defined method, scope, conditions, and result. An item merely absent from the Evidence Basis Record is evidence not identified.
- **Unknown ≠ checked-and-not-observed.** Unknown records that the state cannot be determined from the available basis and procedure. Checked-and-not-observed records only that the specified target was not observed within the defined search or observation conditions.
- **Evidence refused ≠ evidence inaccessible.** Refused records an explicit refusal to provide an item. Inaccessible records inability to access an identified or represented item within the assessment boundary. Neither state establishes what the unavailable item would show.
- **Contradictory ≠ unresolved by default.** Contradictory describes inconsistent evidence relationships. Unresolved/inconclusive describes the proposition-level result after the applicable procedure. Record both where both are true rather than treating them as synonyms.

74. Treatment of Contradictory Evidence

When distinct evidence items or observation paths bear inconsistently on the same proposition, preserve the contradiction and its provenance. Do not resolve it through source preference, evidence class, producer identity, or recency alone unless an identified governing rule or the applicable verification procedure provides a basis for that treatment.

The record should identify:

- the conflicting evidence items or observation paths and their versions/provenance; the proposition and represented condition to which the conflict relates;
- the procedure used to compare or examine the conflict; whether an identified source supersession, version rule, or other governing basis applies;
- what the applicable verification result records about the contradiction; plausible alternative explanations relevant to interpretation; and what remains unresolved.

If resolving the contradiction requires additional evidence, that evidence first enters a revised Evidence Basis Record under §48.1 and is used through an applicable revised verification record. Part IX preserves the contradiction until the controlling verification record changes; it does not silently reconcile the record.

75. Treatment of Unresolved and Inconclusive Results

An unresolved or inconclusive verification result is preserved when the applicable procedure does not resolve the proposition within its stated evidence basis, scope, and conditions.

Record:

- the proposition evaluated; the controlling Evidence Basis Record and verification-record versions; the evidence and procedure actually used;
- the observed or analytical result;
- the specific condition preventing resolution, stated descriptively rather than as an invented adequacy threshold;
- contradictions or plausible alternative explanations relevant to interpretation;
- any evidence-not-identified, inaccessible, refused, or checked-and-not-observed state that contributes to the unresolved result; and what the current record does and does not establish.

A future plan may identify a revised procedure or additional evidence that could bear on the proposition, but that plan is not itself a verification result. Additional evidence follows §48.1, and any revised proposition-level conclusion is recorded through the applicable revised verification record.

Part IX does not convert an unresolved or inconclusive result into Gap, Unknown, Partial, Supported, severity, materiality, or another finding state. Part X forms bounded findings from the preserved verified results and limitations.

76. Worked Example — Handling Negative and Inconclusive Results (Illustrative)

Note: *Illustrative fictional system — system names, record identifiers, configurations, evidence items, procedures, and results below are invented solely to demonstrate the method. They are not ODA3 benchmarks, prescribed test protocols, representations of any real system/provider, or ODA3 assessment outcomes.*

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Negative / Inconclusive Result Record — fictional:

- **Affected proposition:** For the represented retrieval paths in scope, what does the specified evidence support about application of retrieval-time authorization before content is returned?
- **Source verification record:** [fictional Verification Record version/date].
- **Evidence Basis Record:** [fictional Evidence Basis Record version/date].
- **Evidence examined:** [fictional architecture record reference]; [fictional configuration export reference]; [fictional retrieval-log reference].
- **Procedure:** Review the specified records and the controlling verification result for evidence bearing on represented alternate retrieval paths and authorization enforcement within the defined system boundary.
- **State recorded:** Evidence not identified; unresolved/inconclusive verification result.
- **Observation / analytical result:** The stated Evidence Basis Record does not identify an item addressing the represented administrative retrieval path. The specified architecture record does not describe that path, and the retrieval-log evidence used by the controlling verification record does not cover the represented administrative identity context.
- **Limitations / alternative explanations:** This is evidence not identified, not checked-and-not-observed. No defined search or observation procedure establishing absence of the administrative path was performed. The current record therefore does not establish whether such a path exists, whether it invokes the represented authorization mechanism, or whether it can return content.
- **Treatment:** Preserve the unresolved state, source-record references, and limitations for Part X. If additional evidence is later identified, it first enters a revised Evidence Basis Record under §48.1 and

is evaluated through an applicable revised verification record before the Part IX treatment is updated.

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence to preserve: Negative / Inconclusive Result Record version/date; affected proposition; source Verification Record or AI-Specific Behaviour Examination Record version/date; Evidence Basis Record version/date; evidence items or observation paths; procedure/search method where applicable; recorded state or states; observation/analytical result; contradictions; alternative explanations; limitations; and any later revision references.

Conclusion boundary: This record preserves an evidence-not-identified condition and an unresolved/inconclusive proposition-level result concerning the represented administrative retrieval path. It does not establish that such a path exists or does not exist, that authorization is or is not enforced on that path, or that the system is secure or insecure against alternate retrieval paths. Part IX does not assign a finding state. Part X may form a bounded finding only from the controlling verified results and preserved limitations within the defined conclusion boundary.

77. Practitioner Field — Negative and Inconclusive Results Record (Blank)

Use this record to preserve negative, contradictory, checked-and-not-observed, evidence-not-identified, inaccessible, refused, unknown, unresolved, or inconclusive conditions arising from Parts VII and VIII. Select all states that are needed to represent the record accurately.

Assessment Question Reference: _____ Affected Proposition: _____

Evidence Basis Record Version / Date: _____ Source Verification / AI-Specific Behaviour Record Reference / Version / Date: _____ Negative / Inconclusive Result Record Version / Date: _____ State(s) Recorded — select all that apply:

- [ ] Observed negative evidence
- [ ] Contradictory evidence
- [ ] Checked-and-not-observed
- [ ] Evidence not identified
- [ ] Evidence inaccessible
- [ ] Evidence refused
- [ ] Unknown
- [ ] Unresolved / inconclusive verification result Result Context
- Observation / analytical result: _____
- Evidence items / observation paths: _____
- Procedure or search method, where applicable: _____
- Scope, period, conditions, and result: _____
- Provenance / source relationships: _____
- Contradictions and plausible alternative explanations: _____ Limitations and Handoff
- What remains unresolved: _____
- What the current record does not establish: _____
- Relationship to other affected propositions, if any: _____
- Part X handoff reference / preserved limitation: _____ Potential Follow-up (not a verification result)
- Revised procedure that may be considered: _____
- Additional evidence that may be sought: _____
- Evidence Basis Record revision reference, if later created under §48.1: _____
- Revised Verification Record reference, if later created: _____ Resource-planning prompts (completed by the practitioner):
- Internal roles required for follow-up: _____
- External specialist support required: _____
- Estimated practitioner effort: _____
- Access or dependency constraints: _____
- Assumptions and uncertainties: _____

These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration, sample-size, testing, or effort benchmarks.

Part IX Checkpoint

Can a third party who was not involved in the assessment determine:

- which proposition and source verification record the preserved state relates to; which state or states are actually recorded;
- the evidence basis, procedure/search method where applicable, scope, conditions, and result; contradictions, alternative explanations, provenance, and limitations relevant to interpretation;
- what remains unresolved and what the record does not establish; and that any proposed follow-up is distinguished from a completed verification result;

— without the record converting those conditions into a finding state, severity, materiality judgment, conformity decision, certification decision, functional-safety conclusion, or generalized assurance claim?

If not, revise the Negative / Inconclusive Result Record before it is carried into Part X.

The checkpoint is an orientation aid, not a readiness gate, assessment state, certification prerequisite, or ODA3 decision rule.

PART X — FORM FINDINGS AND BOUNDED CONCLUSIONS

78. Why Findings Come After Verification and Negative-Result Treatment

Parts I–IX preserve the assessment question, system boundary, applicability, control understanding, evidence basis, proposition-level verification results, AI-specific examination results, and negative or inconclusive conditions without prematurely converting those records into final conclusions.

Part X forms bounded findings from those controlling records. It does not repair an incomplete evidence basis, perform new verification, or replace the state distinctions preserved in Part IX.

The clean methodological boundary for this Part is:

Part X forms bounded findings from the controlling proposition-level verification results and preserved limitations produced through Parts I–IX. It applies the Supported / Partial / Gap / Unknown descriptive convention defined in §81, used as descriptive finding aids without introducing scoring, weighting, universal thresholds, or proprietary decision logic. Part X does not introduce new evidence, perform additional verification, determine evidence sufficiency by an unstated threshold, assign materiality or severity, or convert a finding into conformity, certification, functional-safety, risk-acceptance, or generalized assurance conclusions. If a finding would require additional evidence or changed verification, the evidence first enters a revised Evidence Basis Record under §48.1 and is evaluated through the applicable revised verification record before the finding is reconsidered.

79. What "Findings" Mean in This Guide

A finding is a bounded synthesis of the applicable control or requirement, controlling verified results, preserved negative/inconclusive conditions, and recorded limitations. It states what those records establish within the defined conclusion boundary.

This guide uses Supported / Partial / Gap / Unknown as descriptive finding aids and does not prescribe a weighted score, numeric threshold, or proprietary finding algorithm. Section 81 therefore defines the bounded public-use semantics for these labels. Those semantics must not be supplemented with unstated evidentiary thresholds, materiality tests, weighting, or implementation criteria.

For each finding, preserve the rationale for the selected state by referencing the controlling records. The state label is not a substitute for the finding statement or conclusion boundary.

A finding applies only to the system state, scope, period, conditions, criteria, evidence basis, and verification records identified in the finding. The same label must not be generalized to unexamined periods, paths, populations, configurations, or criteria.

80. Constructing a Finding

Each finding should contain:

39. **Finding subject and criterion reference:** Identify the applicable control, requirement, or other finding subject established by the frozen assessment basis.
40. **Finding state:** Supported / Partial / Gap / Unknown, using the bounded semantics in §81.
41. **Controlling record references:** Identify the applicable Evidence Basis Record, Verification Record, AI-Specific Behaviour Examination Record, and Negative / Inconclusive Result Record versions where relevant.
42. **Finding rationale:** State how the controlling verified results and preserved conditions bear on the selected finding state without introducing a new verification step.
43. **Contradictory, negative, or inconclusive conditions:** Preserve relevant Part IX states rather than translating them automatically into a particular finding state.
44. **Limitations and alternative explanations:** Record the boundaries that constrain interpretation.
45. **Conclusion boundary:** State exactly what the finding applies to and what it does not establish.

The finding must remain at the scope supported by the controlling records. Where the records do not justify a broader statement, Part X narrows the finding rather than expanding the inference.

81. Applying the Supported / Partial / Gap / Unknown Convention

This guide uses these four labels as descriptive aids. They are not scores, maturity levels, confidence grades, severity levels, or certification states. Apply them to a defined finding subject only after the relevant proposition-level verification results and Part IX conditions have been preserved.

- **Supported:** The controlling verification result supports the defined proposition for the stated scope, period, system state, conditions, evidence basis, and procedure. Recorded limitations remain part of the finding and prevent generalization beyond that boundary.
- **Partial:** The finding subject has been explicitly decomposed into relevant aspects, conditions, paths, or propositions, and the controlling records do not yield one uniform result across them — for example, one bounded aspect is supported while another is contradicted or unresolved. The Finding Record must state the differing components; `Partial` must not be inferred from a percentage, informal notion of “mostly,” or an unstated sufficiency threshold.
- **Gap:** The controlling verification result contradicts the defined proposition within the stated boundary, or verifies within that boundary that a mechanism/state required by the applicable criterion is absent or not implemented. `Gap` is bounded to what was actually verified; it is not automatically created by evidence refusal, inaccessible evidence, evidence not identified, an unresolved contradiction, or a checked-and-not-observed state unless the verification result itself establishes the proposition required for Gap.
- **Unknown:** The controlling records do not resolve the defined proposition within the stated boundary. The Finding Record must preserve why it remains unresolved — for example, evidence not identified, inaccessible or refused evidence, unresolved contradiction, inconclusive verification, or another recorded limitation — without treating that underlying Part IX state as proof of the proposition's truth or falsity.

These labels describe the relationship between the defined finding subject and the controlling assessment record. They do not rank evidentiary strength and do not establish acceptability, materiality, severity, conformity, certification, functional safety, or risk acceptance.

Multiple Part IX states may bear on the same finding. They remain separately visible in the rationale. Part X does not mechanically translate any Part IX state into a finding label.

Where the applicable criterion itself defines a threshold or decision rule, record its authoritative source, version, and scope and apply it only within that scope. COM-020 does not supply a substitute universal threshold.

82. The Boundary Between Finding and Judgment

A finding is not a business, risk-acceptance, legal, regulatory, safety, or certification judgment. The finding records what the controlling assessment records establish within a defined boundary.

For example:

- **Bounded finding statement:** "For the stated production configuration, defined observation period, tested identity/document combinations, and recorded retrieval paths, the controlling verification record supports the proposition that the represented retrieval-time authorization mechanism enforced the examined authorization decisions, subject to the recorded limitations."
- **Generalized judgment:** "The system is protected against unauthorized access."

The second statement extends beyond the bounded verification record unless separately established by an authorized decision process and applicable evidence.

Part X therefore does not convert Supported into "secure," Gap into "unsafe," or any finding state into compliance, acceptability, risk acceptance, certification, or functional-safety status. Governance and other authorized decision functions may use bounded findings as inputs under their own criteria and authority.

83. Worked Example — Forming Findings (Illustrative)

Note: *Illustrative fictional system — system names, record identifiers, criteria, evidence items, procedures, and results below are invented solely to demonstrate method. They are not ODA3 benchmarks, prescribed test protocols, representations of any real system/provider, or ODA3 assessment outcomes.*

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Finding A — represented retrieval-time authorization path

- **Finding subject / criterion reference:** [fictional applicable criterion/control reference].
- **Finding state:** [Supported / Partial / Gap / Unknown — select by applying §81 to the controlling records].
- **Controlling records:** [fictional Evidence Basis Record version/date]; [fictional Verification Record version/date]; [fictional AI-Specific Behaviour Examination Record version/date, if applicable]; [fictional Part IX record version/date, if applicable].
- **Verified result carried forward:** For the tested identity/document combinations, stated observation conditions, and recorded retrieval paths, the controlling verification record states [fictional proposition-level result].
- **Negative / inconclusive conditions:** [fictional preserved Part IX state(s), or "none recorded for this proposition" if that is what the fictional record specifies].
- **Finding rationale:** Apply the applicable §81 definition to the controlling result and preserved conditions. Do not infer operation for untested identities, documents, paths, configurations, periods, or unobserved provider-internal processing.
- **Conclusion boundary:** The selected state applies only to the stated criterion, system configuration, observation period, tested combinations, recorded paths, and controlling record versions. It does not establish generalized security, compliance, certification, or operation outside those boundaries.

Finding B — represented administrative / alternate retrieval path

- **Finding subject / criterion reference:** [fictional applicable criterion/control reference or applicable assessment proposition].
- **Finding state: Unknown** *only if the §81 Unknown definition is satisfied by the fictional controlling record described below.*
- **Controlling records:** [fictional Evidence Basis Record version/date]; [fictional Verification Record version/date]; [fictional Part IX Negative / Inconclusive Result Record version/date].
- **Verified / preserved result:** The Part IX record preserves evidence not identified and an unresolved/inconclusive proposition-level result concerning the represented administrative retrieval path. No defined observation or search establishes absence of that path.
- **Finding rationale:** Under the fictional example, the controlling record does not establish whether the represented alternate path exists, whether authorization is applied on that path, or what behavior would occur through it. The finding state is therefore selected only by applying the §81 Unknown definition to this unresolved record; the Part IX state itself is not treated as an automatic finding-state mapping.
- **Conclusion boundary:** The finding does not establish existence or absence of an alternate path, a bypass, generalized authorization effectiveness, or generalized insecurity.

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence to preserve: Findings Record version/date; finding subject and criterion reference; finding state; §81 finding-state definition applied; controlling Evidence Basis Record, Verification Record, AI-Specific Behaviour Examination Record, and Negative / Inconclusive Result Record references as applicable; verified results carried forward; preserved negative/inconclusive states; rationale; limitations and alternative explanations; conclusion boundary; and revision provenance.

Conclusion boundary: These findings demonstrate how Part X applies the frozen finding-state convention to controlling verified results and preserved limitations for the fictional system. They do not introduce new evidence or verification, establish generalized control effectiveness, legal or regulatory

conformity, certification, functional safety, business acceptability, or risk acceptance. Any change requiring new evidence or verification follows §48.1 and the applicable revised verification record before a finding is reconsidered.

84. Practitioner Field — Findings Record (Blank)

Use this record to form each bounded finding from the controlling verified results and preserved limitations.

Assessment Question Reference: _____ Finding Subject / Control / Requirement Reference: _____ Applicable Criterion / Source / Version: _____

Finding Record Version / Date: _____ §81 Finding-State Definition Applied: _____ Finding State: Supported / Partial / Gap / Unknown **Controlling Record References**

- Evidence Basis Record Version / Date: _____
- Verification Record Reference / Version / Date: _____
- AI-Specific Behaviour Examination Record Reference / Version / Date, if applicable: _____
- Negative / Inconclusive Result Record Reference / Version / Date, if applicable: _____

Verified Results Carried Forward

- Proposition-level result(s): _____
- Observation / analytical result(s) relied upon: _____ Negative / Inconclusive Conditions Preserved
- Applicable Part IX state(s): _____
- Contradictions / alternative explanations: _____
- Evidence not identified / inaccessible / refused, where relevant: _____ Finding Rationale
- How the §81 finding-state definition applies to the controlling record: _____
- Why the selected state is bounded to the stated scope and conditions: _____ Limitations
- Scope / period / configuration / observation limitations: _____
- What remains unresolved: _____
- What the finding does not establish: _____ Conclusion Boundary
- What the finding applies to: _____
- What the finding does not establish: _____ Revision Provenance
- Prior Finding Record reference, if any: _____
- Reason for revision: _____
- Revised Evidence Basis / Verification Record references, if applicable: _____ Resource-planning prompts (completed by the practitioner):
- Internal roles required for finding formation or review: _____
- External specialist support required: _____
- Estimated practitioner effort: _____
- Assumptions and uncertainties: _____

These fields are practitioner-supplied planning information. They are not ODA3 staffing, cost, duration, sample-size, testing, or effort benchmarks.

85. Finding Revision and Handoff

A finding is versioned to the controlling records from which it was formed. It should not be silently overwritten when the evidence basis, verification result, system state, applicable criterion, or conclusion boundary changes.

Where later information requires reconsideration:

46. additional evidence first enters a revised Evidence Basis Record under §48.1; affected propositions are evaluated through the applicable revised verification record;
47. Part IX treatment is updated where negative, contradictory, unresolved, or inconclusive conditions change; and Part X issues a revised Finding Record with explicit provenance to the prior version.

A revised finding does not retroactively change what the earlier finding established for its own recorded basis, scope, period, and conditions.

Part X findings may be handed to governance, risk, compliance, remediation, or other authorized decision functions as bounded inputs. Those functions own their separate decision criteria and authority. Part X does not make those downstream decisions.

Part X Checkpoint

Can a third party who was not involved in the assessment determine:

- the finding subject and applicable criterion; the finding state and the §81 definition used to select it; the controlling record versions and proposition-level results relied upon;
- the negative, contradictory, unresolved, or inconclusive conditions preserved in the rationale; the limitations, alternative explanations, and conclusion boundary;
- whether the finding has been revised and why; and what downstream judgments the finding does not make;

— without the Finding Record introducing new evidence, new verification, an unstated sufficiency/materiality threshold, severity, risk score, conformity decision, certification decision, functional-safety conclusion, or generalized assurance claim?

If not, revise the Finding Record before it is used as an input to downstream governance or remediation activity.

The checkpoint is an orientation aid, not a readiness gate, assessment state, certification prerequisite, or ODA3 decision rule.

PART XI — REMEDIATION AND REASSESSMENT

86. Why Remediation and Reassessment Follow Finding Formation

A bounded finding describes what the controlling assessment records establish for a defined proposition within a stated scope, period, system state, evidence basis, procedure, and conclusion boundary. It does not permanently characterize the system.

After a finding is issued, the assessed system or its surrounding conditions may change. An organization may modify a control, configuration, model, dependency, authorization path, operating procedure, evidence mechanism, human-oversight arrangement, or other element relevant to the finding. Additional evidence may also become available.

Those developments create a change in the assessment state. They do not retrospectively change what the earlier assessment established.

Accordingly, remediation and reassessment must preserve two distinct propositions:

48. What did the previous assessment establish at the time and within the boundary examined?
49. What does a subsequent assessment establish after the relevant change?

The second proposition cannot replace the first. A prior Gap, Partial, Unknown, or Supported finding therefore remains part of the historical assessment record even when later work produces a different finding.

Part XI governs how such changes are routed back through the existing assessment chain. It does not establish a separate remediation-validation methodology, universal reassessment interval, or remediation SLA.

The controlling sequence is:

Historical finding → remediation or change relevant to the frozen assessment boundary, applicability determination, control proposition, evidence basis, or prior finding → boundary/applicability check → revised Evidence Basis Record under §48.1 → revised Part VII verification → revised Part VIII examination where applicable → Part IX treatment where applicable → new versioned Part X finding → preserved relationship to prior finding

A remediation plan, completed engineering task, closed ticket, management representation, newly supplied artifact, or configuration deployment does not by itself alter a finding.

87. Treat Remediation as a Change, Not a Finding Closure

Remediation describes action intended to change a condition associated with a finding. Examples may include changing a configuration, introducing or modifying a control mechanism, altering an authorization path, updating a model or dependency, changing an operating procedure, or modifying an evidence-generation mechanism.

The fact that an action occurred establishes only what the applicable evidence supports about that action.

For example:

- a change ticket may bear on whether a change was authorized or recorded; a deployment record may bear on whether a specified configuration was deployed;
- configuration evidence may bear on the represented implemented state; operating evidence may bear on behavior observed after the change.

None of those evidence classes establishes by classification alone that the earlier finding has changed.

The practitioner should therefore distinguish:

Remediation represented – What action does the organization state was intended or planned?

Remediation implemented – What does the applicable evidence support about the state actually introduced?

Post-change operation examined – What do the applicable verification procedures support about operation under the stated post-change conditions?

Revised finding – What finding follows when §81 is applied to the resulting verified record?

This distinction prevents administrative completion from being treated as assessment closure.

Terms such as closed, resolved, remediated, or fixed should not be used as substitutes for a revised finding unless their meaning is explicitly bounded to an organizational workflow rather than an assessment conclusion.

88. Preserve Historical Findings

A finding is a versioned assessment record.

When later evidence or verification changes the assessment posture, the earlier finding should remain preserved with its original:

- finding identifier and version; proposition or finding subject and §81 finding-state definition applied; assessment boundary;
- applicable system state and period;
- Evidence Basis Record reference/version;
- Verification Record reference/version; relevant Part VIII record reference/version where applicable; relevant Part IX record reference/version where applicable;
- finding state (Supported/Partial/Gap/Unknown per §81); limitations; conclusion boundary; and issuance date or assessment-record timestamp.

A subsequent finding should receive its own version or successor identifier and explicitly reference the prior finding.

The relationship may be recorded as, for example:

Supersedes for the stated post-change assessment state: [prior finding identifier/version]

This means the newer finding controls for the newly examined state. It does not erase or retrospectively invalidate the earlier finding.

Historical preservation is particularly important where a later assessment reaches a different state. A sequence such as Gap → Supported does not mean that the original Gap was erroneous. It may instead mean that the earlier evidence and verification supported a Gap within the earlier boundary and that a later system state, evidence basis, and verification record support a different finding.

The assessment record must preserve enough provenance to distinguish those possibilities.

89. Corrective Evidence Must Re-enter the Assessment Chain

Evidence generated or supplied after remediation does not enter directly into Part X.

It first becomes candidate evidence for the Evidence Basis Record. The §48.1 change discipline therefore continues to apply.

The revised record should identify, as applicable:

- the new or changed evidence item; its source and producer; relevant system state; applicable period; collection or extraction method;
- provenance; relationship to existing evidence; limitations; reason for addition or revision; and resulting Evidence Basis Record version.

Once admitted to the revised Evidence Basis Record, the evidence may be used through the applicable Part VII verification procedure and, where the proposition concerns AI-specific behaviour, the Part VIII specialization of that procedure.

Negative, contradictory, checked-and-not-observed, unresolved, or inconclusive post-change results remain subject to Part IX. Multiple Part IX states may coexist and remain separately visible.

Only after that chain has been completed should Part X be applied again, using §81 definitions directly without relying on any additional decision rules.

Therefore:

New artifact ≠ revised evidence conclusion ≠ revised verification result ≠ revised finding.

Each transition must remain explicit. If the post-remediation examination identifies additional evidence that was not part of the current Evidence Basis Record, that evidence must first enter through §48.1 rather than being informally introduced into the reassessment conclusion.

90. Determine the Reassessment Boundary

Not every system change requires the entire assessment to be repeated.

Part XI does not define a generic "material change" test unless an applicable criterion or the assessment basis already defines materiality. Instead, the practitioner should determine what part of the frozen assessment state may have been affected by a **change relevant to the frozen assessment boundary, applicability determination, control proposition, evidence basis, or prior finding**.

Do not infer relevance merely because a component changed. Ask what the available change information supports about the relationship between the change and the previously assessed propositions.

Relevant change categories may include:

Change dimension	Reassessment question
System boundary	Does the change alter components, actors, dependencies, interfaces, data flows, authority paths, consequence paths, or other elements represented in the frozen boundary?
Applicability	Does the change affect facts used in the existing applicability determination?
Control representation	Does the change alter the represented mechanism, responsibility, dependency, authority, trigger, exception, or operating condition relevant to a proposition?
Evidence basis	Does the change alter the provenance, period, system state, interpretation, or continuing relevance of evidence previously used?
Verification proposition	Does the change affect conditions under which a proposition was previously examined?
AI-specific behaviour	Does the change affect an applicable model, provider, retrieval path, tool path, behavioral dependency, oversight arrangement, or physical-consequence path examined under Part VIII?
Prior finding	Does the change concern a proposition, limitation, contradiction, unresolved condition, or other boundary represented in the prior Finding Record?

The reassessment boundary should then identify which existing propositions require renewed examination and which remain outside the reassessment. A narrow change may justify a narrow reassessment. A broader change may require reopening the system boundary, applicability determination, control understanding, or other upstream records before verification can proceed.

Part XI does not prescribe a universal threshold for making that determination.

91. Point-in-Time, Event-Triggered and Periodic Reassessment

COM-020 distinguishes three reassessment contexts without establishing a universal monitoring or reassessment schedule.

Point-in-time assessment – A point-in-time assessment establishes bounded findings for the system state, period, evidence basis, procedures, and conditions identified in the assessment record. Those findings should not be generalized into a claim of perpetual control operation.

Event-triggered reassessment – A reassessment may follow an identified change relevant to the frozen assessment boundary, applicability determination, control proposition, evidence basis, or prior finding. Examples can include changes to a model, provider, configuration, control mechanism, authorization arrangement, system boundary, evidence mechanism, or other condition relevant to an

assessed proposition. The occurrence of such a change does not automatically determine the reassessment scope. The relationship between the change and the frozen assessment state must first be examined under §90.

Periodic reassessment – An applicable criterion, contractual requirement, organizational methodology, governance arrangement, or defined assessment basis may establish periodic reassessment. Where such a requirement exists, the assessment record should identify its source rather than presenting the interval as an ODA3 universal requirement.

COM-020 does not prescribe:

- a universal reassessment frequency; perpetual or continuous assessment; a remediation deadline; a remediation SLA;
- a universal finding-expiry period; or a universal change threshold or generic materiality definition.

Where continuous observation or monitoring exists operationally, its evidence may bear on relevant propositions. The existence of that monitoring does not convert COM-020 itself into a continuous-assurance methodology.

92. Practitioner Field: Remediation and Reassessment Record

Complete using practitioner-supplied information and references to the controlling assessment records.

Field	Practitioner entry
Prior Finding ID / version	
Prior finding state per §81 and boundary	
Remediation or change reference	
Description of represented change relevant to frozen boundary/applicability/proposition/evidence/finding	
Change evidence reference(s)	
Relevant boundary/applicability/control proposition affected	
Boundary / applicability determination after change	
Basis and limitations for that determination	
Reassessment boundary	
Propositions selected for renewed examination	
Propositions explicitly outside reassessment	
Revised Evidence Basis Record version	
§48.1 revision reference	
Revised Verification Record version(s)	
Revised Part VIII record/version, where applicable	
Revised Part IX record/version, where applicable — state whether multiple states coexist	
New Finding Record ID / version	
§81 Finding-State Definition Applied and how definition applies to controlling revised record	
Relationship to prior finding	
Remaining limitations / unresolved conditions	
Conclusion boundary	

Practitioner checkpoint

Before treating the reassessment as complete, confirm that the record shows:

- what changed and what evidence bears on that change relevant to the frozen boundary/applicability/proposition/evidence/finding;
- which historical finding remains preserved; the boundary/applicability determination made after the change, its basis, and its limitations;
- why the selected reassessment boundary follows from the identified change; which propositions were and were not re-examined;
- that additional evidence entered through the revised Evidence Basis Record under §48.1;
- that applicable propositions were subjected to revised Part VII verification and Part VIII examination where applicable;
- that applicable negative and inconclusive results remain preserved through Part IX with coexisting states where applicable;
- that the resulting finding was formed again under Part X per §81 without relying on any additional decision rules;
- that the new finding identifies its relationship to the historical finding; and that no remediation activity was treated by itself as proof that the previous finding changed.

If these elements are not explicit, revise the Remediation and Reassessment Record and the applicable upstream records before representing a changed assessment finding.

92.1 Worked Example — Reassessing a Finding After Remediation

Fictional scenario: Support Assistant A.

A prior versioned Finding Record contains a Gap concerning a defined authorization proposition for a specified retrieval path and stated assessment conditions.

The organization subsequently represents that the retrieval authorization mechanism has been changed.

Step 1 — Practitioner input

The practitioner records:

- prior Finding Record: [fictional finding ID/version]; remediation/change reference: [fictional change reference];
- represented change relevant to frozen boundary/proposition: [fictional description of authorization-path change];
- affected proposition: [fictional bounded authorization proposition]; affected system state: [fictional post-change system-state identifier];
- candidate post-change evidence: [fictional evidence references]; and proposed reassessment boundary: [fictional bounded scope and conditions].

The practitioner does not mark the prior Gap as closed. The practitioner first records the post-change boundary/applicability determination and the resulting reassessment boundary. Candidate post-change evidence used for renewed examination is then processed through §48.1 into a revised Evidence Basis Record.

Step 2 — Derived assessment output

Using the revised Evidence Basis Record, the applicable proposition is examined again through the Part VII verification discipline and Part VIII specialization where applicable.

Where the revised examination produces negative, contradictory, checked-and-not-observed, unresolved, or inconclusive conditions, those conditions are preserved through Part IX as applicable, with multiple states permitted to coexist.

Part X is then applied to the revised verified record per §81 definitions. Assume, solely for purposes of this fictional example, that the resulting controlling verification record supports the authorization proposition within the stated post-change scope, conditions, system state, evidence basis, and procedure.

A new Finding Record may therefore record Supported for that bounded post-change proposition under §81.

The original Gap Finding Record remains preserved.

Step 3 — Conclusion boundary

Assessment question: For Support Assistant A in the defined production configuration and operating period, what evidence supports the claim that retrieval-time authorization prevents a user from receiving content they would not be permitted to access directly?

Evidence to preserve: Original Finding Record version/date and §81 definition applied; remediation/change reference; revised Evidence Basis Record version via §48.1; revised Verification Record version/date and procedure/scope/period/conditions; revised Part VIII record version where applicable; revised Part IX record version and preserved states; new Finding Record version/date, §81 definition applied and how it applies, linkage to prior finding, limitations, and conclusion boundary.

Conclusion boundary: This record documents what the revised assessment establishes for the specified post-change proposition and boundary. The new finding of Supported applies only to the stated post-change system state, period, evidence basis, and conditions. It does not establish that the earlier Gap never existed, that every authorization path has been remediated, that the mechanism operates identically outside the examined conditions, that all related controls are Supported, that the system is secure, safe, compliant, or certified, or that remediation is permanently effective. The defensible record is not "Gap closed." It is: A historical Gap remains preserved for the earlier assessed state. Following a recorded change relevant to the frozen boundary, revised evidence entry via §48.1, revised verification, Part IX treatment where applicable, a subsequent versioned finding records the proposition state established for the specified post-change assessment boundary per §81.

PART XII — PRACTITIONER FIELD MATERIALS

93. Purpose of Part XII

Parts II through XI each introduce a specific record type at the point in the methodology where it is first needed. Part XII serves as consolidated reference, not repetition of full templates. For each record type introduced in earlier Part, Part XII provides brief purpose, cross-reference to Part where full template and detailed guidance are located, and key fields that must be present. For two new record types (Limitation Register and Claims Checklist), Part XII provides full templates. Part XII also explains how records work together as system, and how version control should be maintained. Where a Part XII summary conflicts with originating frozen Part, originating Part controls; field summaries are non-substitutive.

94. Record Inventory and Cross-References

Record Type	Purpose	Full Template Location
Assessment Question Record	Defines specific question being examined	Part II, §16
System Boundary Record	Identifies components, configurations, data flows, identities, external dependencies within scope	Part III, §24
Applicability Record	Documents which criteria potentially applicable, basis, candidate elements, exclusions, dependencies	Part IV, §33
Control Understanding Record	Documents represented intended control model	Part V, §42
Evidence Basis Record	Identifies, collects, preserves, documents evidence	Part VI, §51
Verification Record	Evaluates defined propositions against versioned Evidence Basis	Part VII, §60
AI-Specific Behaviour Examination Record	Applies Part VII discipline to AI-specific dimensions	Part VIII, §69
Negative and Inconclusive Results Record	Preserves negative, contradictory, checked-and-not-observed, unresolved, inconclusive	Part IX, §77
Findings Record	Forms bounded findings using Supported/Partial/Gap/Unknown per §81 definitions	Part X, §85
Remediation and Reassessment Record	Documents remediation/change and reassessment, preserves historical findings	Part XI, §92
Limitation Register	Consolidates all limitations across assessment, preserves Part IX states	Part XII, §100 (new cross-cutting aid)
Claims Checklist	Supports review of whether proposed statements remain within conclusion boundaries	Part XII, §100 (new cross-cutting aid)

95. Assessment Question Record

Purpose: defines specific question being examined.

Key fields: system/service identity, claim/concern, purpose, consequence if claim false, period/version, environment, authority/independence.

Full template: Part II, §16.

96. System Boundary Record

Purpose: identifies components, configurations, data flows, identities, external dependencies within assessment scope.

Key fields: Assessment Question Reference, System/service, Version/release/configuration, Environment, Operating period, Components material to question (model layer, orchestration layer, agent layer, RAG layer, tools, identities, data flows, monitoring surfaces, provider dependencies, physical consequence path), Explicitly out of scope, Implementation responsibility and evidence-source state.

Full template: Part III, §24. Change affecting or potentially affecting system boundary discovered during examination and effect on existing evidence — neutral provenance, no generic materiality threshold.

97. Applicability Record

Purpose: documents which criteria potentially applicable.

Key fields: Assessment Question Reference, System Boundary Reference, Applicable criteria (source, identifier, authoritative source, disposition, basis), Candidate system element(s), Implementation responsibility and evidence-source state, Dependencies and assumptions, Unresolved interpretation, Exclusions, Regulatory context as input, Prior version, Change affecting or potentially affecting applicability determination identified during examination and effect on existing evidence.

Full template: Part IV, §33.

98. Control Understanding Record

Purpose: documents represented intended control model.

Key fields: Assessment Question Reference, System Boundary Reference, Applicability Record Reference, Candidate System Element, Criterion and intended outcome, Represented mechanism, Ownership and responsibility, Intended operating boundary, Dependencies, Hypothesized failure conditions, Change affecting or potentially affecting represented control understanding identified during examination and effect.

Full template: Part V, §42.

99. Evidence Basis Record

Purpose: identifies, collects, preserves, documents evidence.

Key fields: Assessment Question Reference, System Boundary Reference, Applicability Record Reference, Control Understanding Record Reference, Candidate System Element, Evidence Basis Record Version/Date, Prior Version/Reference, Evidence identification, Evidence provenance, Temporal scope, Coverage description, Known gaps, Inaccessible evidence, Evidence preservation, Change affecting or potentially affecting evidence basis identified during examination, source/reason, dimensions affected, effect on verification work already performed, revised under §48.1.

Full template: Part VI, §51.

100. Verification Record and Related Records — Consolidated View

Verification Record (from Part VII)

Key fields: Assessment Question Reference, System Boundary Reference, Applicability Record Reference, Control Understanding Record Reference, Evidence Basis Record Version/Date, Proposition (ID, statement, subject/mechanism/behavior, scope/path/authority context, relevant period/state, related criterion), Evidence actually used, Verification procedure, Observations and analytical results (observed results, refusals/negative observations, unsuccessful attempts/failures, anomalies, contradictory evidence, checked-and-not-observed state(s) with conditions, unknown/inaccessible/refused evidence), Corroboration and alternative explanations, Proposition-level verification result (bounded result in prose, scope/period/configuration/path, limitations, what result does not establish, unresolved carried forward),

Change affecting or potentially affecting verification result identified during examination, source/reason, propositions affected, prior/current Evidence Basis Record version, procedure/system-state/assumption change, effect on prior result, earlier result preserved, upstream revision required under §58.1. Full template: Part VII, §60.

AI-Specific Behaviour Examination Record (from Part VIII, where applicable)

Assessment Question Reference, Verification Record Reference, Dimension examined (Probabilistic behaviour, RAG behaviour, Agentic systems/tool use, Model/provider change, Physical consequence paths, Human oversight), Proposition examined, Evidence used, Observation/Analytical result, Limitations and alternative explanations, Summary of behavioural limitations. Full template: Part VIII, §69.

Negative / Inconclusive Results Record (from Part IX)

Assessment Question Reference, Affected

Proposition, Evidence Basis Record Version/Date, Source Verification/AI-Specific Behaviour Record Reference/Version/Date, State(s) Recorded (Observed negative evidence, Contradictory evidence, Checked-and-not-observed, Evidence not identified, Evidence inaccessible, Evidence refused, Unknown, Unresolved or inconclusive verification result — multiple states may coexist), Result Context, Limitations and Handoff, Potential Follow-up, Record version/date. Preserves distinct states, does not normalize. Full template: Part IX, §77.

Limitation Register (consolidated; new, cross-cutting aid)

Assessment Question Reference, Limitation Type (Scope exclusion, Evidence gap, Verification constraint, Unresolved contradiction, Inaccessible data, Silent failure risk), Related Record(s) and Version(s), Description, Effect on Conclusion Boundary, Part IX State Preservation Lock: This Limitation Register does not reclassify or replace Part IX states. Where limitation relates to Part IX state, reference Part IX record/state while preserving original classification. Contradictory remains contradictory; checked-and-not-observed retains defined basis; not identified remains distinct from inaccessible/refused; multiple states may coexist, Propositions remaining unresolved within recorded verification boundary, Alternative explanations remaining unresolved or not evaluated within current Evidence Basis Record, Additional evidence or examination identified for possible consideration, if any, without implying resolution, Change in recorded limitation or its relationship to assessment identified during examination.

Finding Record (from Part X)

Assessment Question Reference, Finding Subject/Control/Requirement Reference, Applicable Criterion/Source/Version, Finding State (Supported/Partial/Gap/Unknown per §81 finding-state definitions), §81 Finding-State Definition Applied and how definition applies to controlling record, Controlling Record References, Verified Results Carried Forward, Negative/Inconclusive Conditions Preserved preserving original classification, Finding Rationale (How applicable §81 definition applies to controlling record, why selected state bounded), Limitations, Conclusion Boundary, Prior Finding Record reference, Finding revision provenance. Full template: Part X, §85. Uses §81 directly, not nonexistent mandate rules.

Remediation and Reassessment Record (from Part XI)

Original Finding Reference per §81, Original Evidence Basis/Verification/AI-Specific/Negative-Inconclusive Record Versions, Represented Change, Boundary/Applicability Determination After Change (boundary elements affected/not affected, applicability determination with basis, propositions/findings potentially affected, records examined and limitations), Basis and limitations for that determination, Revised Assessment Chain (Revised Evidence Basis Record version/date via §48.1, Revised Part VII verification version/date and procedure/scope/period/conditions, Revised Part VIII examination where applicable, Revised Negative/Inconclusive Record version/date where applicable — state whether multiple Part IX states coexist), Revised Finding (Revised Finding Record version/date, §81 Finding-State Definition Applied and how applies to controlling revised record, evidence summary, verification summary, limitations and conclusion boundary, linkage to prior finding version and description what changed, finding revision provenance), Reassessment Boundary Determination (findings carried forward unchanged and rationale including basis per §90 — preserve identified change, relationship to boundary/proposition, records examined, limitations; do not infer continued operation beyond boundary — findings requiring reassessment and rationale, findings revised after §89 chain). Canonical lifecycle:

Historical finding → remediation/relevant change → boundary/applicability check → revised Evidence Basis Record (§48.1) → revised Part VII verification → revised Part VIII examination where applicable → Part IX treatment where applicable → new versioned Part X finding per §81 → preserved relationship to prior finding. Full template: Part XI, §92.

Assessment Record Index

Assessment Question ID, System Boundary Record Version/Date, Applicability Record Version/Date, Control Understanding Record Version/Date, Evidence Basis Record Version/Date, Verification Record Version/Date, AI-Specific Behaviour Record Version/Date (if applicable), Negative/Inconclusive Record Version/Date (if applicable), Finding Record Version/Date, Remediation/Reassessment Record Version/Date (if applicable), Limitation Register Version/Date.

Claims Checklist (new, cross-cutting aid — non-authorizing review aid)

Supports review of whether proposed assessment statements remain within recorded finding and conclusion boundaries. Does not authorize, approve, certify, or validate a claim, does not determine publication readiness, does not imply ODA3 authorization. Each checklist item records a review; its determination is recorded separately as Yes / No / Not determined.

- The claim explicitly references the specific Assessment Question ID and system version/environment examined — Outcome: Yes / No / Not determined — Review note: _____
- The claim accurately reflects the Finding State recorded in the Finding Record per §81 definitions, without elevating Partial to Supported or Unknown to Gap without justification — how §81 definition applies: _____ — Outcome: Yes / No / Not determined — Review note: _____
- The claim does not generalize beyond the recorded scope, period, configuration, or observation paths — Outcome: Yes / No / Not determined — Review note: _____
- The claim acknowledges the specific limitations and unresolved conditions recorded in the Finding Record and Limitation Register, preserving Part IX classifications — Outcome: Yes / No / Not determined — Review note: _____
- The claim does not state or imply that the system is "secure," "compliant," "certified," or "audit-ready" in a generalized sense — Outcome: Yes / No / Not determined — Review note: _____
- The claim does not imply that completing this guide constitutes an ODA3 assessment, certification, or conformity determination — Outcome: Yes / No / Not determined — Review note: _____
- The claim states bounded proposition as: "the recorded verification result supports [defined proposition] within [stated boundary/conditions]" — not generic operational claim — Outcome: Yes / No / Not determined — Review note: _____
- Reviewer / function: _____ Review date: _____
- Checklist overall outcome: Yes / No / Not determined — does not authorize, approve, certify, or validate claim — Items recorded as No / Not determined and any resulting follow-up: _____

Part XII Checkpoint

Can a practitioner use these consolidated records to navigate the complete assessment chain from question to finding (Parts II–XI), while preserving distinct states of negative/inconclusive results per Part IX and versioned provenance of all changes, using Limitation Register to locate and cross-reference recorded limitations and Claims Checklist to review proposed claims against controlling records, without introducing new methodological gates, materiality thresholds, or ODA3 benchmarks? If not, consolidation has failed its purpose. Checkpoint is orientation aid, not readiness gate, assessment state, certification prerequisite, or ODA3 decision rule.

CONCLUDING APPARATUS

Source Register / Methodology Notes / Notably Absent / Final Publication Statement

Source Register / Methodology Notes

Source Hierarchy

COM-020 is a practitioner methodology, not a primary source of control requirements. Where an applicable criterion exists, the applicable criterion's authoritative source controls. COM-020 does not create control requirements.

Source and record hierarchy used in this guide (aligned with COM-019):

- **Authoritative criterion source** — the governing framework, regulation, standard, or organizational policy that states the requirement.
- **Represented control model** — the organization's representation of how the requirement is intended to be satisfied, per Control Understanding Record.
- **Implementation responsibility / evidence-source state** — per COM-019 evidence-source-state discipline: who operates/maintains, where evidence originates, and what limitations bound its availability.
- **Evidence Basis material** — evidence identified, collected, preserved, and documented under Part VI, with provenance, temporal scope, coverage, known gaps, inaccessible evidence, and §48.1 change provenance recorded as applicable.
- **Verified proposition result** — proposition-level result from Part VII, specialized by Part VIII where applicable, preserving distinct Part IX states without collapsing into generic limitation.

Framework and Document Relationships

- **COM-019 — The AI System Owner's Guide to Operational Assurance** — COM-020 is the examination perspective corresponding to COM-019's system-owner perspective. COM-020 remains consistent with COM-019's evidence-source-state discipline and Unknown / checked-and-not-observed discipline. COM-019 explicitly states that its practitioner guidance does not create controlled assessment methodology or organizational decision authority; COM-020 preserves that boundary and does not convert guidance into authority.
- **COM-015 — Building an AI Assurance Evidence Pack** — Owns evidence-pack preparation methodology. COM-020 consolidates evidence via Part VI without duplicating COM-015's preparation methodology.
- **COM-017 — Implementing GAISSF™** — Owns implementation guidance for GAISSF™ controls. COM-020 examines represented implementation against defined propositions; it does not prescribe implementation.
- **COM-016** — Not referenced in publication content; relationship to other ODA3 internal work is not exposed per publication boundary.
- **GAISSF™ v1.0, PAI-SF™, NIST AI RMF, ISO/IEC 42001, EU AI Act, OWASP GenAI** — Referenced as potentially applicable criteria sources where organizational applicability determination identifies them; not as authoritative interpretation of those sources.

Methodology Clarifications

- **Bounded findings** — Findings formed per §81 definitions: Supported / Partial / Gap / Unknown — directly applied to controlling verification result within stated scope/period/state/conditions/evidence basis/procedure. No scoring, materiality thresholds, or mandate finding-state decision rules.
- **Part IX preservation** — Negative/inconclusive states (observed negative, contradictory, checked-and-not-observed, evidence not identified, inaccessible, refused, unknown, unresolved/inconclusive) remain distinct and may coexist. Limitation Register observes and references, does not reclassify.

- **Versioning** — Change provenance via §21.1 (boundary), §29.1 (applicability), §40.1 (control understanding), §48.1 (evidence basis), §58.1 (verification), revision provenance (findings), and §92 (remediation/reassessment). Historical findings preserved, not overwritten.
- **Reassessment** — Canonical lifecycle: Historical finding → remediation/relevant change → boundary/applicability check → revised Evidence Basis Record (§48.1) → revised Part VII verification → revised Part VIII examination where applicable → Part IX treatment where applicable → new versioned Part X finding per §81 → preserved relationship to prior finding. No universal continuous-assessment obligation, fixed interval, closure criterion, or remediation SLA.
- **Practitioner field materials** — Part XII consolidates blank templates for the practitioner to complete. No ODA3 staffing, cost, duration, sample-size, testing, or effort benchmarks. Methodological chain Parts II–XI; cross-cutting aids Limitation Register + Claims Checklist observe/reference methodology without altering states.

Source Register (Bibliography)

Publicly available references; practitioner must verify current version and applicability:

- NIST AI Risk Management Framework (AI RMF 1.0) and AI RMF Generative AI Profile
- NIST TEVV Approaches for AI Systems (Initial Public Drafts where applicable)
- ISO/IEC 42001:2023 Artificial Intelligence Management System
- ISO/IEC 23894:2023 AI Risk Management
- EU AI Act (Regulation (EU) 2024/1689) — official text
- OWASP Top 10 for LLM Applications / GenAI
- ODA3 GAISSF™ v1.0 — Global AI Security and Safety Framework (public summary)
- ODA3 PAI-SF™ — Physical AI Security Framework (public summary)
- ODA3 COM-015, COM-017, COM-019 — practitioner guides (public)
- Organizational criteria, contractual requirements, and applicable regulatory guidance as identified via Applicability Record

No proprietary telemetry, internal client data, vendor-specific assessment protocols, or controlled assessment methodology from ODA3 internal work is disclosed in this guide.

Notably Absent — Consolidated Boundary Restatement

Consistent with ODA3 publication discipline, this guide explicitly states what it does not cover, rather than allowing silence to be read as completeness.

This guide does **not**:

- Describe ODA3's own assessment or certification programme, nor grant or imply assessment authority, conformity determination, or certification outcome. Completing this guide's methodology does not produce an ODA3-recognized outcome.
- Establish proprietary scoring, weighted composite maturity, pass/fail thresholds, universal risk scores, or remediation SLAs. Findings are bounded per §81, not scored.
- Create universal materiality thresholds, generic material-change tests, universal reassessment intervals, continuous-assessment obligations, or perpetual monitoring mandates. Reassessment boundaries determined per §90/§91 using "change relevant to..." formulation.
- Provide legal advice, definitive regulatory interpretation (e.g., EU AI Act tier classification, FRIA triggers), financial/investment/TCO advice, or invented cost figures, fixed remediation timelines, or universal benchmarks.
- Provide turnkey implementation code or prescribe single universal preservation control (e.g., specific hashing algorithm) unless required by identified governing source.
- Make detailed functional-safety determinations, grant sector-specific safety approval, or authorize physical return-to-service.
- Convert remediation activity (plan, ticket closure, deployment, assertion, artifact) into automatic Gap closure. Only revised evidence via §48.1, revised verification (Part VII/VIII), Part IX treatment where applicable, and revised finding per §81 changes finding state.
- Overwrite historical findings. Historical findings remain preserved as records of state, scope, period, evidence basis, verification conditions, limitations, conclusion boundary actually examined.
- Collapse distinct Part IX states into generic limitations. Contradictory evidence remains contradictory; checked-and-not-observed retains defined observation/search basis; not identified remains distinct from inaccessible/refused; multiple states may coexist.
- Create parallel assessment methodology via Part XII. Part XII consolidates existing templates and provides two cross-cutting aids (Limitation Register + Claims Checklist) that observe/reference methodology without altering evidence states, verification results, Part IX states, §81 findings, or conclusion boundaries. Where a Part XII summary conflicts with the Part in which the record originates, that Part controls.

Physical AI boundary clarification:

Part VIII explicitly includes physical AI and consequence-path assessment within the scope of this guide. Where the represented boundary and specified evidence concern whether an AI system can initiate or otherwise influence physical action, assessment should identify physical consequence path and hand off detailed physical-AI security treatment to PAI-SF™ and applicable practitioner guidance. The boundary is on safety determination and authority, not on physical AI as a topic. This guide does not authorize physical action, safety case approval, or return-to-service.

Cross-document boundary (COM-019 alignment):

COM-020 is examination perspective corresponding to COM-019 system-owner perspective and remains consistent with COM-019 evidence-source-state and Unknown/checked-and-not-observed discipline. COM-019 explicitly states practitioner guidance does not create controlled assessment methodology or organizational decision authority. COM-020 preserves that boundary: practitioner-owned examination, versioned provenance, bounded findings.

Final Publication Statement

Document ID: ODA3-2026-08-WHP-COM-020 **Title:** How to Assess an AI System — Examination Perspective **Type:** WHP — Adoption Guide (Practitioner Methodology)

Status: Practitioner methodology for organizations to assess their own AI systems against defined control baseline, or to map existing assessment practice to one. Does not describe ODA3's own assessment or certification programme, does not grant or represent assessment authority, does not establish conformity, certification, or independent assurance conclusion. Provides bounded findings per §81 within stated scope/period/state/conditions/evidence basis/procedure, preserving distinct Part IX states and versioned provenance.

Version: Public Version 1.1 — Final Publication. Version 1.0 first published August 2026; Public Version 1.1 issued 28 October 2026.

Structure: Parts I–XII with continuous section numbering §1–§100, followed by unnumbered concluding apparatus.

Licence and attribution

Classification: Public — GAISSF Ecosystem Licence (GEL) v1.0. © 2026 ODA3 Pvt Ltd. Published by ODA3 Institute.

Attribution (GEL v1.0, Section 6): "GAISSF Ecosystem developed by ODA3 Institute, licensed under GEL v1.0." Modified versions must also state: "Modified from GAISSF Ecosystem by ODA3 Institute (ODA3 Pvt Ltd). Original licensed under GEL v1.0."

Commercial use requires a separate written commercial licence from ODA3 Institute. Under GEL v1.0, using this guide as the primary or supporting methodology in any paid engagement is commercial use. ODA3 Institute does not currently offer certification. Licence text: https://oda3.org/gel-v1-0/

Change from Version 1.0

Public Version 1.1 adds the licence, copyright and attribution notices; renames Section 5 to "Recurring Examination Dimensions" and states that those dimensions are not a structure, taxonomy or scoring model; replaces "assessor's companion" and "assessor's perspective" with "examination companion" and "examination perspective"; and adds §21.1 When the Boundary Changes During Examination, which nine places in Version 1.0 already cited but which was missing; numbers the Part XI worked example as §92.1 and renames its steps so they no longer repeat §1–§3; places propositions and template field labels as sub-headings of their sections; and removes production notes. Sections §1–§100 keep their numbers; the method, records and finding states are otherwise unchanged.