ODA3 INSTITUTE

Practitioner Cheat Sheet

AI Security Governance Model

DocID: ODA3-2026-07-CHT-SEC-010

Classification: Public — Practitioner Reference

Version 2.2 | Public | 17 July 2026

Publisher: ODA3 Institute

Publication type: Practitioner Cheat Sheet

Review status: Publication edition

How to use this guide: Establish the governance operating model first; apply the decision framework to individual AI systems; implement the control matrix; connect governance monitoring to UAIF™ incident records and AI-IRF™ response; then test operating effectiveness through the assurance checklist. Adapt thresholds, authorities and retention periods to organizational risk appetite and applicable law.

> **Key definition — AI security governance:** The system of authority, decision rights, accountability, processes, evidence and assurance used to keep AI-related security risk within approved boundaries throughout the AI system lifecycle.

This guide is a practitioner reference, not legal advice or a claim of compliance. Applicability depends on role, jurisdiction, system classification and deployment context.

1. Executive Overview

AI security governance converts principles into enforceable operating decisions. It determines who may approve an AI use case, what evidence must exist before release, which risks may be accepted, who can stop a system, and how security performance is monitored after deployment. Without these mechanisms, an AI policy is descriptive rather than controlling. **[T1/T2 basis: NIST AI RMF 1.0; ISO/IEC 42001:2023 overview and management-system model]**

The risk is not limited to model compromise. AI systems combine models, data, prompts, retrieval stores, tools, identities, third-party services and human decisions. Weak governance permits unregistered systems, uncontrolled data flows, insecure agents, untested model changes, ambiguous incident ownership and risk acceptance by people without authority. The result may be confidentiality loss, unsafe automated actions, service disruption, regulatory exposure or decisions that cannot be reconstructed. **[T2: synthesis of primary standards and documented enforcement cases; likelihood is deployment-specific]**

Governance must connect enterprise oversight with technical gates. The board or governing body sets risk appetite; executives assign accountability and resources; an AI security authority defines minimum controls; system owners own outcomes; engineering teams implement controls; and independent assurance tests whether claimed controls operate. A committee can coordinate these actors, but it does not replace named ownership.

This cheat sheet uses the ODA3 framework suite as its primary operating architecture. **GAISSF™** structures governance, accountability, control objectives, evidence expectations and assurance. **UAIF™** structures the identification, classification, causality and evidentiary record of AI incidents. **AI-IRF™** structures preparation, escalation, containment, eradication, recovery and lessons learned. The loop closes when incident findings and corrective actions return to GAISSF™ governance and control improvement. External frameworks are mapped as independent reference points; those mappings do not establish equivalence, conformity or endorsement.

The immediate objective is a minimum viable governance model: complete inventory, risk-tiered intake, explicit decision rights, security requirements by tier, evidence-backed release approval, continuous monitoring, incident escalation and periodic independent review. CISOs, AI governance leads, security architects, privacy and legal teams, procurement, MLOps, internal audit and business owners must operate one joined process—not parallel review tracks.

> **Notably Absent / Evidence Boundaries:** A governance model does not prove that an AI system is secure, lawful, fair or effective. Those conclusions require scoped technical evidence and applicable legal analysis.

Certification to a management-system standard does not by itself establish the security of every AI system. No evidence reviewed for this edition establishes that the ODA3 suite is the only operational approach, regulator-recognized, accredited, or independently superior. Exact ODA3 control, severity, routing, role and assurance-tier identifiers were not available in the source set for this revision and are therefore not inferred.

Publication scope: This guide does not provide vendor-specific governance tooling recommendations, jurisdiction-by-jurisdiction notification deadlines, legal opinions, or detailed adversarial-ML and red-team procedures. Those matters require controlled, current and topic-specific guidance. Technical content appears only where it enables governance enforcement, evidence, monitoring, containment or assurance.

2. Key Concepts & Definitions

Term	Operational definition
GAISSF™	ODA3 governance and assurance framework used here to organize accountability, control objectives, evidence expectations and continual improvement. Exact internal identifiers require authoritative framework text.
UAIF™	ODA3 incident framework used here to structure AI incident identity, classification, causality and evidence. No severity number is assigned in this edition without authoritative source verification.
AI-IRF™	ODA3 response framework used here to organize preparation, escalation, containment, eradication, recovery and lessons learned for AI incidents.
ODA3 operating loop	GAISSF™ control expectations → UAIF™ incident record → AI-IRF™ response → GAISSF™ corrective action and assurance.
Accountability	Obligation to answer for an AI system's outcomes; it remains with the assigned owner even when tasks are delegated.
AI asset inventory	Controlled record of AI systems, models, datasets, interfaces, owners, suppliers, environments, status and dependencies.
AI security authority	Named function empowered to define security requirements, challenge evidence, approve exceptions and stop unsafe deployment.
AI system owner	Senior person accountable for the system's business purpose, risk treatment, funding and continued operation.
Assurance	Independent, evidence-based evaluation of whether requirements are suitably designed and operating as claimed within scope.
Control objective	Security outcome that must be achieved, independent of the chosen implementation.
Decision right	Explicit authority to approve, reject, pause, restrict or retire an AI system or risk treatment.
Deployer	Organization using an AI system under its authority; the EU AI Act uses this role-specific term.
Exception	Time-bound, approved departure from a control requirement, with compensating measures and expiry.
Foundation model	Broadly capable model that may be adapted or integrated into multiple downstream systems.
General-purpose AI (GPAI) model	EU AI Act term for a model with significant generality capable of competently performing a wide range of distinct tasks.
Governance body	Cross-functional group that coordinates decisions and escalation; it is not the owner of every system risk.
Human oversight	Assigned human capability and authority to understand, supervise, intervene in or stop system operation as required.
Impact assessment	Structured assessment of effects on people, rights, safety, operations and stakeholders over the lifecycle.

Inherent risk	Risk before controls are considered.
Kill switch	Tested mechanism to suspend a model, agent, tool, endpoint or workflow without relying on the affected component.
Model card / system card	Structured record of intended use, limitations, evaluation conditions and performance; not a substitute for deployment-specific testing.
Model risk	Risk arising from incorrect, unstable, manipulated or misapplied model outputs and assumptions.
Operational boundary	Approved users, data, actions, environments, geographies and dependencies within which a system may operate.
Provider	Entity that develops an AI system or GPAI model, or has it developed, and places it on the market or puts it into service under its name or mark.
Residual risk	Risk remaining after implemented controls are considered and validated.
Risk acceptance	Documented decision by an authorized owner to retain a defined residual risk for a stated period.
Security gate	Mandatory lifecycle checkpoint with entry criteria, evidence requirements and approve/reject authority.
Segregation of duties	Separation of build, approval, operation and assurance activities to reduce conflict and single-person failure.
Shadow AI	AI use not recorded or approved through required organizational processes.
Substantial modification	Change that may affect compliance, intended purpose, threat exposure or control effectiveness and therefore triggers reassessment.
System context	People, process, data, infrastructure, models, tools and external dependencies that jointly produce an AI-enabled outcome.
Three lines model	Management owns risk; risk/security functions oversee and challenge; internal audit independently assesses governance and controls.
Evidence tier	Confidence label for a claim: T1 primary verified; T2 corroborated authoritative; T3 limited or single-source; T4 unverified or hypothesis.

3. Threat Landscape

3a. Threat actor taxonomy

Actor type	Motivation	Capability	Likely governance attack or failure path
External attacker	Access, fraud, disruption, influence	Medium–high	Exploit ungoverned endpoint, poison data, inject instructions, compromise supplier or steal model/API credentials.
Malicious insider	Gain, retaliation, concealment	Medium–high	Bypass approval, alter evaluation evidence, export data/model, expand agent permissions or suppress incidents.
Well-intentioned employee	Speed and convenience	Low–medium	Use unsanctioned AI, disclose sensitive data, create unmanaged automations or rely on outputs without validation.
Product or business sponsor	Time-to-market, revenue	Medium	Under-classify risk, narrow test scope, accept risk without authority or describe planned controls as operational.
Supplier / integrator	Commercial delivery	Medium–high	Provide incomplete provenance, change model silently, restrict logs, obscure subprocessors or overstate assurance.
Model or data contributor	Influence or sabotage	Medium–high	Insert backdoor, poisoned sample, unsafe code or misleading documentation into supply chain.

Governance-process attacker	Avoid scrutiny	Medium	Manipulate inventory, evidence, thresholds, metrics or exception records to obtain approval.

3b. Attack vector and path analysis

Governance failures do not map neatly to a single adversarial-technique catalog. MITRE ATLAS references below are representative where a technical attack path exists; “organizational” means no precise ATLAS technique should be claimed.

Vector	Description	Prerequisites	Typical impact	MITRE ATLAS relationship
Shadow AI deployment	System enters use without inventory or review.	Easy SaaS/API access; weak discovery	Data leakage; uncontrolled decisions	Reconnaissance/resource development may precede use; governance failure itself is organizational.
Risk-tier manipulation	Scope or autonomy is understated to obtain lighter review.	Self-attestation; no challenge	Inadequate testing and oversight	Organizational; may enable later ML attack staging.
Evidence tampering	Evaluation, logs or approvals are edited, omitted or selectively presented.	Mutable evidence store; conflicted reviewer	False assurance; poor incident reconstruction	Related to impairing defenses; validate against current ATLAS release before formal mapping.
Supplier opacity	Model/data/version changes occur without notice or evidence.	Weak contract and telemetry	Unknown exposure; broken controls	ML supply-chain compromise family.
Privilege expansion	Agent gains new tools, credentials or execution scope without reassessment.	Weak change control	Unauthorized action; lateral movement	Abuse of ML-enabled product; conventional identity TTPs also apply.
Exception accumulation	Temporary waivers become permanent operating state.	No expiry/owner/monitoring	Systemic control debt	Organizational.
Metric gaming	Teams optimize approval metrics rather than real security outcomes.	Fixed thresholds; no adversarial review	Hidden failure modes	Organizational.
Committee diffusion	Collective review leaves no one accountable for action.	Ambiguous RACI	Delayed containment; accepted unknown risk	Organizational.

3c. Impact analysis

Impact category	Example	Cheat-sheet priority	UAIF™ treatment	Stakeholders
Technical	Agent uses over-privileged tool to alter production data.	Critical	Create incident identity; classify impact and affected system; preserve action evidence	Engineering, customers, CISO
Financial	Fraudulent output triggers payment or contractual decision.	High–critical	Record financial consequence, causality status and decision trace	CFO, business owner, customers
Regulatory	High-risk system lacks required risk management, logs or oversight.	High	Record applicable role, obligation context and evidentiary gaps without making a legal conclusion	Legal, compliance, governing body
Reputational	Organization makes unsupported security or AI capability claims.	High	Link claim, approval record, evidence and affected stakeholders	Board, communications,

				customers
Safety	Model output influences clinical, industrial or transport action outside validated boundary.	Critical	Prioritize safety consequence, exposure and causal uncertainty	Users, safety officers, regulators
Operational	Supplier model update breaks evaluation assumptions and service controls.	High	Record component/version change and resulting control failure	Operations, product, customers

The priority labels above are operational triage labels for this cheat sheet, not asserted UAIF™ severity-tier identifiers.

3d. Documented governance cases

Case	Evidence	Governance failure / root cause	Defender lesson	Notably absent
FTC v. Rite Aid facial-recognition matter (2023 order)	**T1:** regulator complaint/order	The FTC alleged failure to use reasonable procedures and prevent consumer harm from false-positive matches; testing and oversight were inadequate for the use context.	Require deployment-specific performance evidence, documented response to harmful outcomes and executive ownership.	The matter does not prove that all facial-recognition deployments create the same risks or legal outcome.
SEC “AI washing” enforcement actions (2024)	**T1:** regulator orders and release	The SEC found that two investment advisers made false and misleading statements about their purported use of AI.	Require accountable approval and traceable substantiation for public AI and security claims.	The actions do not establish a universal disclosure standard for every AI statement or jurisdiction.
Moffatt v. Air Canada (British Columbia CRT, 2024)	**T1:** tribunal decision	The tribunal held Air Canada responsible for inaccurate information provided by its website chatbot in the dispute before it.	Assign an owner for consequential external outputs; test, monitor and preserve a human resolution path.	The decision does not establish that every chatbot error creates liability; outcome depends on facts and applicable law.
Samsung source-code exposure via public LLM (2023)	**T2:** multiple contemporaneous news reports; corroborated by subsequent corporate policy change	Employees pasted proprietary source code into a public generative-AI tool with no enterprise gateway, inventory or egress control in place.	Pair acceptable-use policy with technical enforcement (gateway routing, egress controls); policy alone does not prevent shadow AI.	No verified evidence of subsequent third-party exploitation of the exposed code.
Clearview AI biometric-data enforcement actions (2020–2024)	**T2:** multiple European data-protection authority fines (Italy, UK, and others), publicly reported	Biometric data was processed at scale with no accountable governance review of lawful basis, purpose limitation or privacy-by-design at the point the system was built.	Governance review must occur at design time, not be retrofitted after deployment; privacy and legal review are governance functions, not afterthoughts.	Enforcement outcomes and appeals differ by jurisdiction; this does not establish a uniform global legal standard.

For each case, GAISSF™ provides the governance/control lens, UAIF™ provides the incident identity and evidence lens, and AI-IRF™ provides the response and corrective-action lens. The case records do not by themselves validate exact ODA3 identifiers or prove control effectiveness.

4. Technical Deep Dive

4.1 Operating architecture

```
GAISSF™ — establish ownership, boundaries, control objectives and evidence
    ↓ expected state and assurance criteria
Runtime — AI system operates, changes and produces telemetry
    ↓ event or suspected harm
UAIF™ — establish incident identity, classification, causality status and evidence
    ↓ routing basis and structured incident record
AI-IRF™ — prepare, identify, contain, eradicate, recover and learn
    ↓ validated corrective actions and residual-risk decision
GAISSF™ — update controls, assurance scope, risk treatment and oversight
```

This loop sits across the organizational layers below. GAISSF™ is not limited to the executive layer; UAIF™ is not only a reporting label; and AI-IRF™ is not activated only after technical compromise. Each supplies a distinct operational function.

Layer	Components	Required governance output
Enterprise direction	Governing body, executive risk committee, risk appetite	GAISSF™: risk appetite, prohibited uses and delegated authorities
Portfolio control	AI governance office, CISO, privacy, legal, procurement	GAISSF™: inventory, control baseline, exception register and exposure
System accountability	Business owner, product owner, system owner	GAISSF™: intended use, boundary, assessments and risk decision
Engineering control	Security architecture, data, MLOps, platform, red team	GAISSF™ evidence; UAIF™-ready telemetry; AI-IRF™ containment capability
Operations	SOC, model monitoring, SRE, incident response	UAIF™ incident record; AI-IRF™ response, rollback and recovery evidence
Independent assurance	Compliance, second-line validation, internal audit	GAISSF™: challenge, sampling, findings and improvement verification

Control flow:

```
Business proposal → inventory/intake → inherent-risk tier
        ↓                  ↓                ↓
  named owner       applicable duties   control baseline
        ↓                  ↓                ↓
threat + impact assessment → build/test evidence → release authority
                                                        ↓
                              monitoring → change/incident trigger
                                                        ↓
                                    reassess, restrict or retire
```

4.2 Decision rights and segregation of duties

RACI entries below are a baseline. Every organization must resolve them to named roles and delegated-authority thresholds.

Decision	Governing body	Syste	AI	CISO /	Legal /	Engineerin	Independent

	/ executive	m owne r	governance / risk	security	privacy	g / MLOps	assurance
Set AI risk appetite and prohibited uses	**A**	I	R	C	C	I	C
Approve intended use and fund controls	I	**A/R**	C	C	C	C	I
Assign risk tier and control baseline	I	C	**A/R**	C	C	C	I
Design and implement technical controls	I	A	C	C	C	**R**	I
Approve release within delegated authority	I	**A**	R	C	C	C	I
Accept residual Critical risk or prohibited-use exception	**A**	R	C	C	C	I	I
Stop or restrict an unsafe system	I	A	C	**R**	C	R	I
Determine legal/regulatory notification	I	C	C	C	**A/R**	I	I
Verify control operation independently	I	I	C	C	C	I	**A/R**

R = Responsible; A = Accountable; C = Consulted; I = Informed. A governance committee may coordinate a decision but must not obscure the single accountable authority.

4.3 Governance structure options

The structure may vary; the required outcomes do not. Every option must preserve named system ownership, independent challenge, consistent minimum controls, escalation authority and enterprise-level visibility.

Structure	Operating model	Suitable context	Primary risk	Required safeguard
Cross-functional AI governance committee	Existing functions jointly review defined decisions under an executive charter	Organizations with a limited or moderately sized AI portfolio	Committee becomes advisory and no one owns the decision	Named accountable chair, quorum, delegated authority and decision log
Central AI governance office	Dedicated team owns methods, portfolio oversight, control baselines and coordination	Large or highly regulated organizations requiring consistency	Bottleneck, distance from engineering and excessive centralization	Risk-tiered delegation, service-level objectives and embedded technical partners
Embedded governance within enterprise GRC/security	AI governance uses established risk, compliance and security processes	Organizations with mature GRC and ISMS capabilities	AI-specific threats, evidence and lifecycle changes are diluted	AI-specific competence, controls, telemetry and assurance procedures
Federated governance	Business units own systems and first-line decisions; a central function sets minimum rules and challenges high risk	Decentralized or multi-product enterprises	Inconsistent tiering, exceptions and evidence across units	Common taxonomy, central inventory, mandatory gates and portfolio assurance
Hybrid model	Central policy, assurance and high-risk authority with embedded business/engineering governance	Complex organizations balancing consistency and delivery speed	Authority boundaries become ambiguous	Published decision-rights matrix and one escalation path per decision class

Do not select a structure solely from organization size. Base the decision on portfolio diversity, consequence, regulatory exposure, engineering decentralization, assurance capacity and the speed at which material changes occur.

4.4 Policy framework and enforceable gates

Policy family	Minimum enforceable rule	Gate or evidence
Acceptable AI use	Define approved services, prohibited data and prohibited actions	Gateway policy, endpoint controls, user acknowledgement
AI system intake	No real data, external user or production action before registry and owner assignment	Registry ID required in deployment metadata
Risk and impact	Review depth follows consequence, autonomy, data, exposure and legal context	Completed assessment and challenge record
Secure development	Threat model and security tests attach to the exact system version	Signed evidence bundle in release pipeline
Human oversight	Consequential actions require competent, informed and timely intervention capability	Tested approval/override path
Supplier governance	Model and service changes, logs, provenance and security duties are contractually addressed	Due-diligence file and change notification
Monitoring and incidents	Runtime must emit telemetry sufficient for UAIF™ classification and AI-IRF™ response	Logging test and alert-to-ticket trace
Exceptions	Every deviation has an owner, compensating control, expiry and authorized acceptance	Exception record checked automatically at release
Retirement	Access, endpoints, retained data and residual obligations are closed	Decommissioning evidence and registry status

Conceptual release policy:

```
DENY production release when any condition is true:
  registry.system_id is missing
  accountable_owner is missing
  approval.expired = true
  runtime_boundary != approved_boundary
  unresolved_finding.severity = Critical
  required_evidence_bundle.signature_valid != true
  exception.expired = true

REQUIRE enhanced approval when:
  sensitive_data OR external_users OR consequential_decision
  OR privileged_tool_access OR material_supplier_change
```

The policy logic is illustrative. It must be translated into organization-specific thresholds and tested against emergency-change and rollback procedures.

4.5 How governance bypass occurs

1. A team begins with an apparently low-risk experiment.
2. Production data, external users, retrieval, tools or autonomy are added incrementally.
3. No event is classified as a material change, so the original approval remains.
4. Logging and evaluation remain calibrated to the prototype boundary.
5. A model or supplier update changes behaviour without a new security decision.
6. An incident occurs, but ownership, rollback authority and evidence are unclear.
7. The committee retrospectively discovers that the system operating in production is not the system it approved.

4.6 Exploit and failure conditions

Condition	Diagnostic test
Incomplete inventory	Compare expense, DNS, proxy, API-key, cloud marketplace, code and

	identity records with registry.
Self-selected risk tier	Require second-person challenge for external-facing, sensitive-data, safety or agentic use.
No system boundary	Ask for one current diagram showing model, data, tools, identities and decision consumers.
Approval without evidence	Sample approvals and locate immutable test artefacts linked to exact version/configuration.
Uncontrolled supplier change	Verify contract notice, version pinning, release notes and regression triggers.
Risk acceptance without authority	Match signatory authority to residual-risk rating and financial/safety exposure.
Policy–operation gap	Trace one requirement from policy to control, telemetry, owner and test result.

4.7 Technical indicators

Stage	Observable indicator
Discovery	AI API domains, packages, model downloads or OAuth grants absent from inventory.
Boundary expansion	New tool, dataset, user group, geography, action or privilege without assessment ticket.
Gate bypass	Production deployment lacks signed release record or evidence hashes.
Control drift	Model/configuration identifier differs from approved bill of materials.
Exception abuse	Expired exception remains open; compensating-control telemetry is missing.
Incident suppression	Security event exists in technical logs but not AI incident register.

4.8 Hardened pattern versus vulnerable baseline

Vulnerable baseline	Hardened pattern
Annual policy attestation	Event-driven lifecycle gates plus scheduled review
Spreadsheet inventory maintained manually	Registry reconciled against technical discovery sources
Committee owns “AI risk”	Named accountable system owner; committee coordinates/challenges
Vendor questionnaire accepted as proof	Contractual rights plus version-specific evidence and testing
Generic pentest	AI-specific threat model, misuse cases and adversarial evaluation
Approval attaches to product name	Approval binds model, configuration, data, tools, identity and use boundary
Logs for troubleshooting only	Security telemetry supports detection, reconstruction and oversight
Risk waiver without end date	Time-bound exception with owner, compensating controls and expiry

4.9 Minimum machine-readable governance record

```
{
  "system_id": "ai-000123",
  "owner": "role:vp-customer-operations",
  "intended_use": "draft customer support responses",
  "prohibited_uses": ["automatic refunds", "legal commitments"],
  "risk_tier": "high",
  "components": {
    "model": "provider/model@immutable-version",
    "retrieval_index": "kb-prod@2026-07-10",
    "tools": ["ticket.read", "draft.write"],
    "data_classes": ["internal", "customer-confidential"]
  },
  "approvals": ["security", "privacy", "business-owner"],
  "evidence_bundle": "sha256:...",
```

```
  "approved_until": "2026-10-17",
  "kill_switch_owner": "role:on-call-service-owner"
}
```

Example governance drift query:

```
SELECT runtime.system_id, runtime.model_version, registry.approved_model_version
FROM runtime_inventory runtime
JOIN ai_registry registry USING (system_id)
WHERE runtime.model_version <> registry.approved_model_version
   OR registry.approved_until < CURRENT_TIMESTAMP;
```

5. Practitioner Decision Framework

5a. Risk scoring matrix

This is an ODA3 practitioner scoring aid for prioritization, not a verified UAIF™ severity method or regulatory classification. Validate weights against organizational loss scenarios and review performance after incidents. Score each factor 1–4, multiply by weight, and divide the total by 100. Any hard-stop condition overrides the numeric score.

Risk factor	Weight	1 — Low	2 — Medium	3 — High	4 — Critical
Decision consequence	20	Reversible draft	Limited internal effect	Material customer/employee effect	Safety, rights or major financial effect
Autonomy/tool authority	15	No action	Human-approved action	Bounded automated action	Irreversible or privileged action
Data sensitivity	15	Public	Internal	Confidential/personal	Special-category, secrets or regulated
Exposure	10	Isolated test	Internal users	External bounded users	Internet-scale or critical service
Model/supply-chain control	10	Fully controlled/pinned	Managed supplier	Limited transparency	Mutable/unknown provenance
Detectability	10	Full trace and replay	Strong logs	Partial reconstruction	Material actions unobservable
Security novelty	10	Established pattern	Minor novelty	Agent/RAG/multimodal complexity	Novel architecture with no test baseline
Legal/regulatory relevance	10	None identified	Contractual	Regulated context	High-risk/prohibited-use concern

Tier thresholds: 1.00–1.49 Low; 1.50–2.24 Moderate; 2.25–2.99 High; 3.00–4.00 Critical.

Hard stops: no accountable owner; prohibited use suspected; unapproved sensitive data; no shutdown path for consequential autonomy; critical evidence missing; unresolved critical security finding.

5b. Decision tree

```
Is the AI use recorded with a named accountable owner?
  NO  → Do not deploy; register or stop the use.
  YES → Is a prohibited-use or legal hard-stop concern present?
        YES → Escalate to Legal/Compliance and AI Security Authority; hold deployment.
        NO  → Does it process sensitive data, affect people materially, or execute actions?
              YES → High/Critical review: threat model, impact assessment,
                    adversarial testing, executive risk decision and monitored release.
              NO  → Is it externally exposed or dependent on a mutable supplier?
                    YES → Moderate review plus supplier/change controls.
                    NO  → Low-tier baseline, inventory and periodic verification.
```

```
After approval: Did model, data, tools, autonomy, users, geography or purpose change?
  YES → Treat as reassessment trigger before continued operation.
  NO  → Continue monitoring until scheduled review or incident trigger.
```

5c. Prioritization

Scenario	Risk	Action	Owner	Timeframe
Unregistered external AI endpoint	Critical	Restrict access; identify owner; preserve logs; assess	CISO + business executive	Immediate / 24 h
Agent has production write access without transaction limits	Critical	Disable write tool; review identity and actions	Service owner	Immediate
Approved model version changed by supplier	High	Pin/rollback; regression test; reassess	Product + MLOps	24–72 h
Expired exception	High	Remove exception or stop affected function	Risk owner	5 business days
Inventory lacks technical reconciliation	Medium	Deploy discovery and monthly reconciliation	AI governance office	30 days
Low-risk internal drafting assistant	Low	Baseline controls, user rules and quarterly sampling	Business owner	Before use

5d. Delegated risk authority

Decision class	Minimum authority	Mandatory consultation	Non-delegable condition
Low residual risk within approved boundary	System owner	Security where control changed	Prohibited use or missing owner
Moderate residual risk or time-bound exception	Business executive / designated risk owner	AI governance and security	Exception exceeds delegated duration or exposure
High residual risk, consequential external use or material supplier opacity	Executive risk authority	CISO, legal/privacy and affected business owner	Governing-body threshold reached
Critical residual risk, safety/rights exposure or prohibited-use question	Governing body or expressly delegated executive committee	CISO, legal, safety and independent challenge	Required legal prohibition; missing essential evidence

Authority must be specified in policy and tied to measurable thresholds. A system owner cannot accept risk on behalf of affected people, regulators, customers or another accountable executive merely because the technical team can deploy the system.

6. Security Controls Matrix

Mappings are functional crosswalks, not declarations of conformity. Obtain licensed standards text before audit use. “EU AI Act” applies only where the relevant role and system classification bring the requirement into scope.

Control	Implementable requirement	GAISSF™	UAIF™	AI-IRF™	Evidence expected	Type; eff./compl	External alignment

						ex.	
GOV-01 Inventory reconciliation	Reconcile registry monthly against proxy, cloud, code, identity, expense and API-key sources.	Portfolio governance and asset accountability	Supplies system identity if an incident occurs	Supplies owner/dependency context for response	Reconciliation query, mismatches, closure tickets	P/D; H/M	NIST AI RMF GOVERN/MAP; ISO 27001 A.5.9; CSF ID.AM
GOV-02 Named accountability	Assign an executive system owner and operational owner before processing real data.	Accountability and decision rights	Identifies accountable incident stakeholders	Defines escalation and recovery authority	Approved RACI, authority record	P; H/L	NIST AI RMF GOVERN; ISO 42001 leadership/roles; CSF GV.RR
GOV-03 Risk-tiered intake	Score consequence, autonomy, data, exposure, supply chain and detectability; independently challenge high tiers.	Risk governance and control selection	Establishes context for later incident classification	Sets response readiness proportional to exposure	Intake, challenge, approval	P; H/M	NIST AI RMF MAP; ISO 42001 risk process; EU AI Act Art. 9 where applicable
GOV-04 Approved system boundary	Version model, data, retrieval, prompts, tools, identities, users and prohibited uses.	Control scope and assurance boundary	Defines affected components and evidence scope	Enables isolation and recovery scoping	Diagram, manifest, hashes	P; H/M	NIST AI RMF MAP; ISO 27001 A.8.27; EU AI Act Arts. 9, 11, 13 context
GOV-05 Evidence-bound release	Bind approval to immutable component versions and signed test/evidence bundle.	Evidence expectation and release assurance	Preserves pre-incident baseline	Enables comparison, rollback and recovery validation	Evidence index, signatures, decision	P; H/H	NIST AI RMF MEASURE/MANAGE; ISO 42001 operation/evaluation; CSF PR.PS
GOV-06 Supplier governance	Contract for provenance, security notice, logs, subprocessors, version change and testing rights.	Third-party control and accountability	Records supplier/component involvement	Defines supplier escalation and joint response	Contract, due diligence, notices	P/D; H/M	NIST AI RMF GOVERN; CSF GV.SC; ISO 27001 A.5.19–23
GOV-07 Least-authority agents	Separate agent identities; allow-list tools/actions; set transaction/time limits; require approval for high-impact actions.	Control objective for autonomy and access	Captures unauthorized actions and affected resources	Supports isolation, revocation and containment	IAM state, action policy, tests	P/D; H/H	OWASP LLM06 (2025); ISO 27001 access controls; EU AI Act Arts. 14–15 where applicable
GOV-08 Material-change triggers	Reassess changes to model, prompt policy, data, retrieval, tools, autonomy, users, geography or purpose.	Lifecycle governance and reapproval	Links incidents to change history	Triggers preparedness and rollback review	Change diff, assessment, approval	P; H/M	NIST AI RMF MANAGE; ISO 42001 change/operation; EU substantial-modification context
GOV-09 Governance telemetry	Alert on unregistered endpoints, version drift, expired approvals, privilege growth and missing logs.	Control-performance monitoring	Provides time, identity and classification evidence	Activates identification and containment	Rules, test events, response tickets	D; H/H	NIST AI RMF MEASURE; CSF DE.CM/DE.AE; EU AI Act logging duties where applicable
GOV-10	Pre-authorize stop,	Governance	Requires	Directly	Runbook,	P/	NIST AI RMF

Incident authority	isolate, rollback, evidence preservation and legal/regulatory escalation.	authority and risk treatment	structured incident identity and impact record	supports all response phases	contact tree, exercise	C; H/M	MANAGE; CSF RS/RC; ISO 27001 A.5.24–28
GOV-11 Exception governance	Require owner, rationale, residual risk, compensating controls, expiry and independent challenge.	Exception and residual-risk governance	Records exception contribution to incident causality	Guides containment and corrective action	Exception register and telemetry	P/D; M/L	NIST AI RMF GOVERN; ISO management-system documented information
GOV-12 Independent assurance	Test risk-based samples and trace claims to live evidence; report findings to oversight.	Assurance and continual improvement	Tests incident records for completeness and consistency	Tests response, recovery and lessons closure	Workpapers, findings, retest	D/C; H/M	NIST AI RMF GOVERN/MEASURE; ISO 42001 internal audit/review; CSF GV.OV
GOV-13 Claims substantiation	Require evidence-linked approval for security, compliance and AI-capability claims.	Evidence discipline and oversight	Preserves claim-related incident evidence	Supports correction and disclosure decisions	Claim, source, approver, limitations	P/D; H/L	NIST AI RMF GOVERN; ISO documented information; EU transparency context
GOV-14 Decommissioning	Revoke identities, remove endpoints, archive required evidence, address retained data/models and update inventory.	Lifecycle closure and residual obligations	Records post-service incidents against retained assets	Defines shutdown, archive and residual monitoring	Closure checklist and access proof	P/C; M/M	NIST AI RMF MANAGE; ISO 27001 A.8.10; CSF ID.AM/PR.DS

Mapping boundary: ODA3 entries above are functional mappings derived from the verified framework purposes available for this task; they are not exact control-clause mappings. External alignments are functional crosswalks, not equivalence, compliance, certification or endorsement determinations.

7. Detection & Monitoring

Monitoring serves three linked purposes: GAISSF™ tests whether the approved control environment remains in force; UAIF™ receives the identity, impact, causality and evidence needed to create a defensible incident record; AI-IRF™ receives the telemetry needed to identify, contain and recover. A dashboard that cannot support these decisions is operational reporting, not sufficient governance evidence.

7a. Logging requirements

Log source	Required fields	Minimum retention basis	Priority
AI gateway/model API	system, user/service ID, model/version, prompt/output reference or protected digest, policy result, tokens, latency, tool calls, trace ID	Risk/legal schedule; high-risk AI Act logs where applicable	P1
Identity and secrets	principal, role, grant/revoke, token scope, source, MFA, timestamp	Security retention schedule	P1
Agent/tool runtime	requested action, arguments classification, authorization, approval, result, rollback reference	Sufficient to reconstruct consequential actions	P1
Registry/approval	owner, tier, boundary, approver, evidence hash, expiry, exceptions	Lifecycle plus required assurance period	P1

MLOps/ deployment	model/data/prompt/config hashes, environment, deployer, pipeline, test result	Lifecycle plus rollback window	P1
Data/retrieval	source, classification, ingest/change, access, index version, deletion	Privacy, contractual and incident needs	P1
Supplier	release/version, service notice, outage/security notice, subprocessor change	Contract and assurance schedule	P2
User feedback	system/version, issue type, consequence, escalation, disposition	Risk-based	P2

Retention is not “keep everything.” Minimize personal and confidential content, separate content from security metadata, restrict access, and document legal holds and deletion.

7b. Detection logic

```
Rule: Unregistered AI Runtime
Source: DNS/proxy/API gateway/cloud inventory
Logic: observed_ai_service = true AND system_id NOT IN approved_registry
Threshold: Any production or sensitive-data event
False-positive risk: Medium
Tuning: Maintain sanctioned experimentation ranges; require owner tags.
```

```
Rule: Approved Boundary Drift
Source: Deployment + registry + agent runtime
Logic: runtime(model,data,tools,privileges,users,region) != approved_boundary
Threshold: One material mismatch
False-positive risk: Low–Medium
Tuning: Normalize immutable identifiers; distinguish emergency rollback.
```

```
Rule: Governance Exception Failure
Source: Exception register + control telemetry
Logic: expiry < now OR compensating_control_status != healthy
Threshold: Any High/Critical system; daily aggregation otherwise
False-positive risk: Low
Tuning: Suppression requires a new approved exception, not ticket extension.
```

```
Rule: Risk-Tier Evasion Signal
Source: Intake + architecture metadata
Logic: declared_tier <= Moderate AND
       (external_users OR sensitive_data OR write_tool OR consequential_decision)
Threshold: Any match before release
False-positive risk: Medium
Tuning: Use as review trigger, not automatic accusation.
```

7c. Indicators of governance compromise

Indicator	Confidence	Notes
Production system absent from registry	High	Verify discovery accuracy and ownership immediately.
Approval record created after deployment	High	Possible retrospective approval or migration issue.
Evidence hash does not match stored artefact	High	Treat as evidence-integrity incident.
Repeated tier downgrade immediately before gate	Medium	Needs independent review of rationale.
Shared administrator identity performs build and approval	High	Segregation-of-duties failure.

Runtime version cannot be determined	High	Approval boundary cannot be verified.
Supplier change without notice	Medium –high	Validate contract, version and impact.

7d. Telemetry pipeline

Feed AI gateway, identity, MLOps, data lineage, agent runtime, supplier and registry events into a normalized schema keyed by `system_id`, `component_id`, `version`, `principal`, `trace_id` and timestamp. Correlate live state with the approved boundary. Send high-confidence boundary violations to the SOC; send portfolio trends and overdue remediation to governance oversight. Protect governance telemetry as a high-value security system and monitor administrator access to it.

7e. Governance dashboard

Targets must be approved from organizational risk appetite and tested baseline data; the table deliberately does not prescribe universal percentages.

Metric	Definition	Decision supported	Recipient
Inventory reconciliation coverage	Discovered AI assets matched to an approved registry record / AI assets discovered	Whether shadow-AI discovery and ownership are effective	AI governance lead; CISO
Boundary-conformance rate	Production systems matching approved model, data, tool, privilege and region boundary / systems tested	Whether approvals remain valid	System owners; security architecture
Evidence-complete release rate	Releases with all mandatory signed evidence / releases requiring a gate	Whether policy is enforced in delivery	Engineering leadership; assurance
Open critical findings	Critical findings not closed or accepted by authorized authority	Whether deployment or continued operation should stop	CISO; executive risk authority
Exception exposure	Open exceptions by risk, age, affected systems and compensating-control health	Whether temporary deviations are becoming normal state	Risk committee
Time to governance decision	Time from validated trigger to documented authorized decision	Whether escalation authority works under pressure	CISO; governing oversight
Incident loop-closure rate	Closed AI incidents with verified corrective action and GAISSF™ control reassessment / closed AI incidents	Whether UAIF™ and AI-IRF™ findings improve governance	AI governance; internal audit
Supplier-change reassessment rate	Material supplier changes reassessed before continued use / material changes received	Whether third-party change control operates	Procurement; system owner
Assurance finding recurrence	Repeated findings with same causal condition	Whether remediation addresses root cause	Audit committee; management review

Report counts and denominators together. A “100% complete” metric is unreliable when the population is unknown, the denominator excludes shadow AI, or teams can self-mark controls complete without evidence.

8. Incident Response Guidance

Use UAIF™ to establish and maintain the incident record before and during response. Use AI-IRF™ to govern the response phases below. Feed validated lessons, residual-risk decisions and corrective actions back into

GAISSF™ assurance and control improvement. Phase names below are functional descriptions; exact AI-IRF™ internal identifiers were not verified from source documents for this edition.

Preparation

Step	Action	Owner	Evidence	Decision point
1	Define AI incident taxonomy and severity	CISO	Approved taxonomy	Which events invoke executive/legal review?
2	Pre-authorize stop, isolate, model rollback and tool revocation	System owner	Authority matrix; test record	Who can act outside business hours?
3	Exercise supplier and agent incident scenarios	IR lead	Exercise report	Are contractual/logging gaps acceptable?

Identification

Step	Action	Owner	Evidence	Decision point
1	Confirm system, version, boundary and owner	SOC	Registry + runtime state	Is the system ungoverned or drifted?
2	Classify technical, rights, safety and regulatory impact	IR lead	Timeline; affected decisions	High/Critical escalation?
3	Preserve prompts/outputs, model/config hashes, tool actions, identities and approvals	Forensics	Evidence manifest; hashes	Legal hold or privilege required?

Containment

Step	Action	Owner	Evidence	Decision point
1	Suspend affected endpoint/tool/identity or switch to human-only workflow	Service owner	Change record	Full shutdown or bounded restriction?
2	Pin or roll back model, prompt policy, retrieval index and configuration as a unit	MLOps	Before/after hashes	Is prior version demonstrably safe for this case?
3	Block propagation to downstream automations	Architect	Dependency graph	Which decisions need review or reversal?

Eradication

Step	Action	Owner	Evidence	Decision point
1	Remove compromised data, prompt, dependency, credential or permission	Engineering	Remediation diff	Has root cause been removed?
2	Correct governance failure: owner, gate, exception, supplier or monitoring	Governance lead	Updated control record	Is this systemic across portfolio?
3	Re-test abuse and regression cases	Independent tester	Signed report	Residual risk within authority?

Recovery

Step	Action	Owner	Evidence	Decision point
1	Restore in stages with enhanced monitoring	Service owner	Release approval	Proceed, restrict or roll back?
2	Review affected historical decisions/actions	Business owner	Review sample and remediation	Notify or compensate affected parties?
3	Confirm owner accepts scoped residual risk	Executive owner	Signed risk decision	Approval expiry and conditions?

Lessons learned

Step	Action	Owner	Evidence	Decision point
1	Complete root-cause analysis across technical and governance layers	IR lead	Causal analysis	Which controls failed or were absent?
2	Search portfolio for the same condition	AI security authority	Hunt results	Enterprise-wide stop/change needed?
3	Report facts, uncertainties and notably absent evidence	CISO	Final report	External disclosure/notification?

Mandatory loop closure: convert confirmed root causes into GAISSF™ control changes, assign corrective-action owners and deadlines, update the assurance plan, and verify operating effectiveness. Record the final incident classification and evidence disposition using UAIF™. Do not close the AI-IRF™ response solely because service has resumed.

Internal escalation template

```
System / owner:
Observed event and time:
Confirmed facts:
Potential consequences:
Containment completed:
Evidence preserved:
Unknowns / notably absent:
Decision required, authority and deadline:
```

External holding statement template

```
We identified an issue affecting [bounded service/function] on [date/time].
We [suspended/restricted] the affected function and are investigating.
Confirmed impact: [facts only]. We have not confirmed: [material unknowns].
We will provide updates through [channel].
```

Notification duties vary by facts, jurisdiction and organizational role. Legal counsel should assess AI-specific, cybersecurity, privacy, sectoral, contractual and safety triggers; do not assume one regime displaces another.

9. Audit & Assurance Checklist

9a. Documentation requirements

Artefact	Purpose	Owner	Review
AI governance charter	Authority, scope and	Governing body	Annual / material change

	escalation	delegate	
Risk appetite and prohibited uses	Decision boundaries	Governing body	Annual
AI inventory	Portfolio completeness	AI governance office	Continuous; monthly reconcile
Responsibility matrix	Named accountability	Executive sponsor	Quarterly
Risk-tiering method	Consistent review depth	CISO / enterprise risk	Annual
System boundary record	Approval scope	System owner	Every material change
Threat and impact assessments	Risk identification	Security + governance	Pre-release / change
Control and evidence matrix	Trace requirements to proof	Security architect	Each release
Exception register	Govern deviations	Risk function	Monthly
Incident/monitoring plan	Operational readiness	IR/service owner	Semiannual exercise
Supplier assurance file	Third-party dependency	Procurement	Annual/change
Management review pack	Performance and decisions	AI governance lead	Quarterly

9b. Technical evidence

Control area	Evidence	Collection	Assessor expectation
Inventory	Discovery-to-register reconciliation	Re-run query and sample mismatches	Completeness measured, not asserted
Boundary	Versioned architecture and runtime manifest	Compare live and approved states	Exact components and privileges match
Testing	Raw results, cases, environment, hashes	Reproduce sample	Failures and limitations retained
Access	Identity and tool grants	Export authoritative IAM state	Least privilege and separation operate
Monitoring	Rules, events, response tickets	Trigger controlled test	Detection produces timely action
Change	Deployment records and reassessments	Sample material changes	Triggers consistently applied
Exceptions	Waivers and compensating telemetry	Sample open/expired items	Authority, expiry and monitoring evident
Incidents	Logs, timeline, decisions, lessons	Trace one incident end-to-end	Facts, unknowns and remediation preserved

9c. Assurance interview questions

Ask these ten questions across the Security Architect, MLOps Engineer, AI Governance Lead and CISO; require examples and evidence rather than yes/no answers.

8. Show the authoritative boundary of one production AI system and prove the runtime matches it.
9. Who can accept its residual security risk, and what limits that authority?
10. Which changes trigger reassessment, and show the latest example.
11. How do you discover AI uses that teams did not register?
12. What would cause immediate shutdown, and who can execute it now?
13. How are model, data, retrieval, prompt policy, tools and identities versioned together?
14. Show a failed evaluation and how the organization responded.
15. Which supplier facts cannot currently be verified, and how is that uncertainty treated?
16. Demonstrate one governance monitoring rule from event through closure.

17. What evidence supports management's latest claim about AI security posture, and what is notably absent?

9d. Maturity indicators

Level	Observable characteristics	Evidence
1 — Ad hoc	Policy fragments; voluntary review; incomplete ownership	Inconsistent records; retrospective approvals
2 — Defined	Charter, inventory, tiers and baseline controls exist	Templates and assigned roles; uneven operation
3 — Operational	Gates, monitoring, incidents and exceptions work across portfolio	Samples show repeatable decisions and timely action
4 — Measured	Coverage, drift, control performance and remediation are quantified	Trend data; threshold rationale; management decisions
5 — Adaptive	Threat, incident and assurance evidence changes controls and risk appetite	Portfolio-wide learning; predictive triggers; tested resilience

10. Research Insight ★

Research Insight

Finding: NIST, *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile* (NIST AI 600-1), July 2024. **[T1 — primary government publication]**

Summary: The profile identifies governance as a cross-cutting response to generative-AI risks and emphasizes clear roles, inventories, third-party considerations, incident disclosure, testing and ongoing monitoring. Its structure shows that model evaluation alone cannot manage risks arising from organizational use, supply chains and changing deployment context.

Technical significance: Security properties depend on the complete deployed system and its operational boundary, not only the base model. Governance must therefore bind approval and evidence to model version, data, retrieval, tools, identities, users and intended use.

Validation of guidance: The profile supports GOV-01 (inventory), GOV-04 (security boundary), GOV-06 (supplier control), GOV-08 (change triggers), GOV-09 (governance telemetry) and GOV-10 (incident authority). These controls operationalize the ODA3 loop by establishing GAISSF™ governance evidence, generating UAIF™-usable incident context and enabling AI-IRF™ response decisions.

Operational lesson: Tomorrow, select one consequential AI system and compare its live components and permissions to its approved record. Treat every unexplained mismatch as a governance finding.

Notably Absent: NIST AI 600-1 does not validate ODA3 framework conformance, exact ODA3 identifiers, comparative uniqueness or the effectiveness of every control in every deployment. The mapping above is an evidence-bounded operational interpretation.

Primary reference: [NIST AI 600-1, Generative AI Profile](https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf).

11. Common Mistakes

Mistake	Why it matters / correct approach
Treating an AI council as the risk owner	Committees coordinate; name one accountable system owner and specific decision authorities.
Starting with policy but no inventory	An unknown system cannot be governed; reconcile registry with technical discovery.
Governing the model instead of the system	Risks arise from data, tools, users, identities and workflow; approve a bounded configuration.
Using one control baseline for every use	Apply risk-tiered depth with hard stops for sensitive, consequential and autonomous uses.
Accepting vendor certification as deployment proof	Supplier assurance does not validate your integration, data, permissions or use context.
Allowing system owners to select and approve their own tier	Require independent challenge where consequence or ambiguity is material.
Treating human oversight as a user near the loop	Define competence, information, time, authority and tested intervention mechanism.
Recording only successful tests	Preserve failures, limitations, overrides and excluded cases; otherwise assurance is biased.
Approving a product name without version binding	Tie approval to immutable model, configuration, data, tools and evidence identifiers.
Permitting indefinite exceptions	Require expiry, compensating telemetry, risk authority and reassessment.
Monitoring model quality but not governance drift	Detect unregistered use, version mismatch, privilege expansion and expired approval.
Assuming AI incidents belong only to data science	Integrate SOC, IR, privacy, legal, safety, supplier and business escalation.
Reporting only what happened	State material unknowns and what notably did not occur or was not evidenced.
Claiming compliance from a mapping	A crosswalk supports analysis; it is not a conformity determination.

12. Quick Reference Tables

12.0 ODA3 operating-loop reference

Operational question	Primary framework function	Required output	Feeds next
What must be governed and evidenced?	GAISSF™	Owner, boundary, control objective, evidence and risk decision	Runtime monitoring and UAIF™ context
What happened, how is it classified and what evidence supports that view?	UAIF™	Stable incident identity, classification, impact, causality status and evidence record	AI-IRF™ routing and response
What must be contained, removed, restored and learned?	AI-IRF™	Response decisions, preserved evidence, recovery validation and corrective actions	GAISSF™ improvement and assurance
Did corrective action change the control environment effectively?	GAISSF™	Updated control, reassessment, assurance test and residual-risk decision	Continued monitoring

12.1 Failure-to-control mapping

Failure/attack	Primary controls
Shadow AI	GOV-01 inventory reconciliation; GOV-09 telemetry
Risk-tier manipulation	GOV-03 tiered intake; GOV-12 independent assurance
Evidence tampering	GOV-05 signed evidence gate; access segregation
Supplier opacity	GOV-06 supplier control; GOV-08 change triggers
Agent privilege expansion	GOV-07 least authority; GOV-09 drift detection
Exception accumulation	GOV-11 exception governance
Metric gaming	GOV-12 independent assurance; outcome sampling
Committee diffusion	GOV-02 named accountability; authority matrix

12.2 Risk-to-control mapping

Risk	Controls
Sensitive data disclosure	Boundary, data classification, gateway policy, supplier terms, monitoring
Unauthorized agent action	Separate identity, allow-listed tools, transaction limits, human approval, kill switch
Unsafe model update	Version pinning, change notice, regression testing, staged release, rollback
Unsubstantiated assurance claim	Evidence-linked claims approval, independent review, limitations statement
Unreconstructable decision	Trace IDs, versioned components, protected logs, retention schedule
Regulatory role confusion	Role/classification assessment, legal review, obligation register

12.3 Security implementation checklist

#	Check	Yes	No	Partial
1	Every AI system has a stable ID and accountable owner.	☐	☐	☐
2	Registry is reconciled with technical discovery.	☐	☐	☐
3	Intended and prohibited uses are documented.	☐	☐	☐
4	Risk tier is independently challenged where required.	☐	☐	☐
5	System boundary includes data, model, retrieval, tools and identities.	☐	☐	☐
6	Threat and impact assessments reflect deployment context.	☐	☐	☐
7	Controls are traced to version-specific evidence.	☐	☐	☐
8	Critical tests have no unresolved critical finding.	☐	☐	☐
9	Agent permissions use least authority and action limits.	☐	☐	☐
10	Supplier changes trigger notice and reassessment.	☐	☐	☐
11	Runtime drift from approval is detected.	☐	☐	☐
12	Logs reconstruct consequential actions.	☐	☐	☐
13	Shutdown and rollback are tested.	☐	☐	☐
14	Exceptions expire and compensating controls are monitored.	☐	☐	☐
15	Decommissioning revokes access and closes data obligations.	☐	☐	☐

12.4 Governance readiness checklist

#	Governance check	Status
1	Governing body approved AI risk appetite and prohibited uses.	☐
2	Decision rights and escalation authorities are explicit.	☐
3	Business, security, privacy, legal, safety and audit roles are integrated.	☐

4	Provider/deployer and regulatory classifications are recorded.	☐
5	Portfolio metrics report coverage and control operation, not activity alone.	☐
6	Management reviews open critical risk and exceptions.	☐
7	Claims about AI/security require evidence and approval.	☐
8	Incident reporting includes uncertainties and notably absent evidence.	☐
9	Competence requirements exist for approvers and operators.	☐
10	Independent assurance has access and reporting authority.	☐

12.5 Audit readiness checklist

#	Evidence check	Status
1	Current governance charter and approval history	☐
2	Complete inventory with reconciliation results	☐
3	Sampled system boundaries and runtime manifests	☐
4	Risk-tier rationale and independent challenge	☐
5	Requirements-to-control-to-evidence traceability	☐
6	Raw evaluation evidence including failures	☐
7	Supplier due diligence and change records	☐
8	Open/expired exception and remediation records	☐
9	Monitoring alerts and incident response evidence	☐
10	Management review decisions and follow-through	☐

13. Certification Readiness Notes

This section supports assessment readiness; it does not state that ODA3 certification, accreditation, licensed assessment delivery, or regulatory recognition is currently operational. Exact GAISSF™ assurance tiers and scheme rules must be taken from authoritative current ODA3 publications before use in an engagement.

Framework / standard	What assessors look for	Common nonconformity	Evidence	Maturity indicator
NIST AI RMF 1.0	GOVERN embedded across MAP, MEASURE and MANAGE; risk treatment is contextual and iterative	Framework categories copied into policy without operating process	Profiles, inventory, assessments, measures, decisions	Risk evidence changes priorities and controls
ISO/IEC 42001:2023	AIMS scope, leadership, risk/opportunity process, operational controls, evaluation and improvement	Scope excludes material AI activity; Annex A treatment lacks rationale	Scope, objectives, SoA/treatment, audit and management review	Repeatable AIMS linked to actual AI systems
ISO/IEC 27001:2022	AI security integrated with ISMS assets, suppliers, access, development, logging and incidents	AI treated outside ISMS despite shared information risks	Risk treatment, controls, internal audit	Unified cyber/AI control operation
NIST CSF 2.0	GOVERN informs Identify, Protect, Detect, Respond and Recover	AI profile has no measurable current/target state	Organizational profile, actions and metrics	Portfolio drift and response are measured

ISO/IEC 38507:2022	Governing body evaluates, directs and monitors organizational use of AI	Operational committee substitutes for governing-body direction	Principles, delegated authorities, oversight reporting	Board-level decisions tied to evidence
ISO/IEC 42005:2025	Impact assessment process is scoped, documented and lifecycle-aware	Generic checklist with no affected-party or context analysis	Impact assessment record and actions	Impact evidence triggers design/change decisions
EU AI Act	Correct role/classification and applicable requirements supported by technical/QMS records	Organization assumes all AI has identical duties or uses policy as proof	Risk, data, technical documentation, logs, oversight, monitoring	Obligations integrated with release and operations
ODA3 framework suite: GAISSF™ / UAIF™ / AI-IRF™	Governance controls connect to structured incident classification, response evidence and continual improvement	Framework names appear in policy but incident records, response actions and assurance evidence are disconnected	GAISSF™ control/evidence mapping; UAIF™-structured incident records; AI-IRF™ response and loop-closure evidence	One traceable operating loop from governance decision through incident learning

Examiner traps

18. **Inventory theatre:** a polished list exists, but discovery sources show unregistered systems and no completeness metric.
19. **Human-oversight theatre:** a human is nominally present but lacks time, information, competence or authority to intervene.
20. **Evidence theatre:** dashboards and certifications are presented without version, scope, raw results, exceptions or proof of operating effectiveness.

Tier 1 versus Tier 2 assessment

Dimension	Tier 1 — documentation review	Tier 2 — technical and operating test
Purpose	Determine whether governance design and required artefacts exist	Determine whether controls operate and withstand challenge
Sampling	Policies, records, selected systems and interviews	Live systems, configurations, telemetry, replay and adversarial tests
Inventory	Review stated registry and method	Reconcile against independent discovery
Approval	Inspect workflow and sign-offs	Trace approval to exact runtime state and reproduce sample evidence
Monitoring	Review rules and dashboards	Generate controlled events and verify detection/response
Conclusion	Design readiness within reviewed scope	Operating effectiveness for tested systems and period

Recommended preparation timeline

Period	Objective
Weeks 1–2	Confirm scope, roles, authority, regulatory applicability and evidence repository.

Weeks 3–6	Reconcile inventory; tier systems; identify high-risk gaps and hard stops.
Weeks 7–12	Implement lifecycle gates, control baselines, supplier requirements, telemetry and incident authority.
Weeks 13–16	Run internal operating tests; close critical findings; exercise shutdown and incident paths.
Weeks 17–20	Complete independent readiness review, management review and evidence remediation.

Timeline depends on portfolio size and existing control maturity. Do not schedule certification before inventory completeness and high-risk operating evidence can be demonstrated.

Primary References

- [NIST AI Risk Management Framework 1.0](https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf)
- [NIST AI 600-1: Generative Artificial Intelligence Profile](https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf)
- [NIST Cybersecurity Framework 2.0](https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf)
- [NIST SP 800-61 Rev. 3: Incident Response Recommendations and Considerations for Cybersecurity Risk Management](https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-61r3.pdf)
- [ISO/IEC 42001:2023 overview](https://www.iso.org/standard/81230.html)
- [ISO/IEC 42005:2025 overview](https://www.iso.org/standard/44545.html)
- [ISO/IEC 42006:2025 overview](https://www.iso.org/standard/42006)
- [ISO/IEC 38507:2022 overview](https://www.iso.org/standard/56641.html)
- [Regulation (EU) 2024/1689 — Artificial Intelligence Act](https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng)
- [FTC Rite Aid facial-recognition enforcement](https://www.ftc.gov/news-events/news/press-releases/2023/12/rite-aid-banned-using-ai-facial-recognition-after-ftc-says-retailer-deployed-technology-without)
- [SEC enforcement actions concerning misleading AI claims](https://www.sec.gov/newsroom/press-releases/2024-36)
- [Moffatt v. Air Canada, 2024 BCCRT 149](https://www.canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html)
- [MITRE ATLAS](https://atlas.mitre.org/)
- [OWASP Top 10 for LLM Applications 2025](https://genai.owasp.org/llm-top-10/)


```
Doc ID: CHT-SEC-010 (pending confirmation against file listing)
Classification: Public
Intended Audience: CISOs, Security Architects, AI Governance Leads, Compliance Officers,
Standards Participants, Auditors, Board and Risk Committee Members
Evidence Confidence: T1–T2 for cited primary authorities and documented cases;
T3 where an operational synthesis is not independently validated
Review Cycle: Quarterly recommended
Framework Cross-References: GAISSF™ v1.0; UAIF™ v1.0; AI-IRF™ v1.0;
NIST AI RMF 1.0; NIST CSF 2.0; ISO/IEC 42001:2023;
ISO/IEC 42005:2025; ISO/IEC 42006:2025; ISO/IEC 38507:2022;
ISO/IEC 27001:2022; OWASP Top 10 for LLM Applications; MITRE ATLAS;
EU AI Act
Source Cut-off Date: 17 July 2026
Known Evidence Gaps: Exact ODA3 control, severity, routing, role, evidence-field and
assurance-tier identifiers were not verified from authoritative framework documents for
this revision. External standards mappings are functional and require licensed source text
before audit-grade clause use. Samsung and Clearview AI case entries are T2 (corroborated
public reporting), not primary regulatory/tribunal documents.
Related Publications: CHT-SEC-009 (AI Security Evidence Collection Guide)
```

GAISSF™, UAIF™, and AI-IRF™ are frameworks of ODA3 Institute, referenced under the GAISSF Ecosystem License (GEL) v1.0.

Publisher: ODA3 Institute