Apptega

GUIDE

ISO/IEC 42001:2023 (AIMS)

Compliance Guide

Current as of October 2, 2025

Apptega

Purpose

Convert ISO/IEC 42001:2023 requirements into a practical, auditable **AI Management System (AIMS)** your organization can operate and defend during certification.
This guide specifies owners, organization-defined parameters (ODPs), measurable acceptance criteria, evidence, continuous monitoring, internal audit, management review, and corrective actions. Each clause (4–10) follows a consistent pattern: **Intent → Minimums → Implement → Evidence → Acceptance Criteria → Common Failures → Internal QA → Docs**, with applied narratives and auditor-ready packs.

How to use this guide

1. Stand up the core AIMS with the Quick-Start (appoint roles, set risk appetite, define scope and authorization boundary, start the risk and evidence registers).
2. Implement requirements by clause in the Standards & Practices section using the common pattern above.
3. Set thresholds in **ODPs** (e.g., TEVV pass/fail, incident SLAs, logging baseline, access controls) and file artifacts to the **Evidence Register** as you go.
4. Run the governance loop on cadence: **Continuous Monitoring → Internal Audit → Management Review → Corrective Action → Change Management**.
5. Use the annexes for boundary templates, TEVV plans, supplier responsibility matrices, auditor packs, and communications.

Apptega

Apptega

1. Introduction

ISO/IEC 42001:2023 sets out the requirements for establishing, implementing, maintaining, and continually improving an AI Management System (AIMS). This guide translates those requirements into an operating program your organization can adopt, measure, and defend during certification. It is ISO/IEC 42001–specific and uses the standard's structure (Clauses 4–10) while adding practical detail for AI lifecycle execution — risk-based testing and evaluation (TEVV), data rights and lineage, human oversight, transparency and records, incident response, and supplier governance.

Apptega

What this guide delivers

- A program blueprint for an AIMS aligned to Clauses 4–10, with owners, **organization-defined parameters (ODPs)**, acceptance criteria, and auditable evidence.
- Operational procedures for the AI lifecycle (data, models, prompts/policies, guardrails, deployment, monitoring, and change) that meet 42001's intent without over-prescribing technology.
- A governance loop that links day-to-day AI operations to leadership oversight: **continuous monitoring, internal audit, management review, and corrective action**, with clear triggers and timelines.
- Ready-to-use artifacts: scope and authorization boundary templates, risk and obligations registers, TEVV plans, supplier responsibility matrices, incident/AAR templates, audit programs, and management review packs.

Who should use it

AIMS Executive Sponsor, AIMS Manager, Risk Owner(s), AI Lifecycle Owners (data, model, deployment), Security/Privacy, Legal/IP & Ethics, Supplier/TPRM, SRE/Model Ops, Internal Audit, and Business Owners accountable for AI use cases.

How to work with it

Start with the Quick-Start to appoint roles, set risk appetite, define the **AIMS scope and authorization boundary**, and stand up core registers (risk and evidence). Use the Standards & Practices sections to implement requirements by clause with the pattern Intent → Minimums → Implement → Evidence → Acceptance Criteria → Common Failures → Internal QA → Docs: Configure **ODPs** (thresholds for TEVV, logging, incident SLAs, access controls) and file supporting artifacts to the **Evidence Register** as you go. Operate the control loop on the defined cadence: monitor KRIs, audit, review with leadership, and close corrective actions with verified effectiveness.

What “good” looks like

- **Scope & context are explicit:** use cases, intended use, constraints, data categories/sources, deployment channels, supplier list, and change authorities are documented and under change control.
- **Risks are owned and measurable:** risk criteria and appetite are approved; Tiered use cases drive TEVV depth; residual risk is within appetite or formally accepted with expiry and compensating controls.
- **TEVV is risk-tiered and repeatable:** safety, robustness/jailbreak, factuality, privacy/IP duplication, and bias/fairness evaluations have thresholds, pass/fail records, and reruns on change.
- **Human oversight and transparency are real:** reviewer criteria, escalation thresholds, and user disclosures exist, are tested, and generate evidence.
- **Suppliers are governed:** responsibility matrices, due diligence, incident/change notices, and telemetry/evaluation rights are current and enforced.
- **Records prove the system works:** logs, evaluations, incidents/AARs, management reviews, internal audits, and corrective actions show a feedback loop that reduces recurrence.

1A. Beginner Quick-Start (First 30—90 Days)

Days 1—15: Stand up AIMS

- Appoint AIMS Executive Sponsor, AIMS Manager, Risk Owner(s), AI Lifecycle Owners (data, model, deployment), Security/Privacy Lead, Legal/IP & Ethics, TPRM Lead, Internal Audit Lead.
- Approve AIMS Policy and risk appetite (safety, reliability, privacy, fairness, IP).
- Define AIMS scope & authorization boundary; start Risk Register & Evidence Register; adopt document control.

Days 16—45: Implement Core

- Context & stakeholder analysis; obligations register.
- Risk method (likelihood/impact) and treatment workflow.
- AIMS objectives with measurable targets and KPI/KRIs.
- Draft TEVV plan (safety, factuality, robustness/jailbreak, bias/fairness, privacy, IP/duplication).
- Baseline change control; incident taxonomy & severity; supplier responsibility matrices.

Days 46—90: Validate & Evidence

- Limited-scope internal audit; management review.
- One AI incident/rollback drill; one evaluation rerun on a critical use case.
- Record nonconformities and corrective actions with effectiveness checks.
- Publish AIMS Manual (v1); finalize certification prep plan.

Quick-Start Checklist

Task	Owner	Evidence	Due-by
Appoint core roles	Exec Sponsor	Role matrix; announcement	Day 7
Approve Policy & Risk Appetite	Exec Sponsor	Signed policy; minutes	Day 10
Define scope & boundary	AIMS Manager	Scope statement; diagrams	Day 15
Risk & Evidence Registers live	Risk Owner / AIMS Mgr	CSV/Repo	Day 15
Draft TEVV plan	TEVV Lead	TEVV v1	Day 30
Set ODP thresholds/KRIs	AIMS Mgr / Risk Owner	ODP sheet	Day 30
Change control & incident taxonomy	AIMS Mgr / Sec	SOPs	Day 30
Limited internal audit & MR	Audit Lead / Exec	Audit rpt; MR minutes	Day 75–90
Publish AIMS Manual v1	AIMS Manager	Manual v1	Day 90

2. AIMS Scope & Alignment

- **Scope:** All material AI use cases, datasets, prompts/policies, models, pipelines (training/fine-tune/RAG/inference), safety layers, tools/functions, telemetry, and external providers affecting declared outcomes.
- **Exclusions:** Written justification + authorized risk acceptance; annual re-validation.
- **Risk tiering:** T1 Baseline • T2 Moderate • T3 High.
- **Traceability:** For each use case: tier, owner, intended use & constraints, data sources, deployment channels, suppliers, change authority; boundary artifacts are versioned and change-controlled.
- **Acceptance:** Scope approved; tiered; exclusions justified/time-boxed; flows/diagrams current; suppliers mapped to responsibility matrices.

Apptega

3. Roles & Responsibilities (RACI)

Activity / Artifact	Exec Sponsor	AIMS Mgr	Risk Owner(s)	Data Owner	Model Owner	Sec/ Privacy	Legal/IP & Ethics	Procure-ment/ TPRM	SRE/ Model Ops	Internal Audit	Backup/ Delegate
AIMS Policy & Risk Appetite	A	R	C	C	C	C	C	I	I	I	Named per role
Scope & Authorization Boundary	A	R	C	C	C	C	C	I	C	I	Named per role
Risk Method & Register	C	R	A	C	C	C	I	I	I	C	Named per role
TEVV Plan & Execution	I	R	C	C	A	C	I	I	C	I	Named per role
Change Control (Tiered Gates)	I	R	C	C	A	C	I	I	C	I	Named per role
Incident Mgmt & AARs	I	R	C	C	C	A	I	I	C	I	Named per role
Supplier Due Diligence & Matrices	I	C	I	I	I	C	C	R/A	I	I	Named per role
Internal Audit Program	I	C	I	I	I	C	I	I	I	R/A	Named per role
Management Review	A	R	C	I	I	C	I	I	I	C	Named per role
Corrective Actions & Effectiveness	A	R	C	C	C	C	I	I	C	C	Named per role

Verification note: Exec Sponsor signs MR minutes & risk acceptances; AIMS Manager maintains Evidence Register & ODPs; Internal Audit issues CAPA reports with effectiveness verification.

A = Accountable, R = Responsible, C = Consulted, I = Informed.

Apptega

4. 42001 Clauses (4—10) at a Glance — How to use: run a quick gap scan with C/PC/NC/N/A

Clause	Theme	Outcomes When "Good"
4	Context & Scope	Stakeholders & obligations known; perimeter & data flows documented and change-controlled.
5	Leadership	Policy, objectives, roles, resources, cadence; decisions & exceptions tracked to expiry.
6	Planning	Risk criteria & treatment; measurable objectives; AI change planning integrated.
7	Support	Competence verified; comms planned; document control enforced; sensitive docs access-controlled.
8	Operation	AI lifecycle planned/controlled; tiered TEVV; incidents & suppliers managed with evidence.
9	Performance Evaluation	KRIs monitored; internal audits executed; MR drives actions to closure.
10	Improvement	RCA/CAPA verified effective; continual improvement demonstrable.

Apptega

5. Standards & Practices (Deep Dives by Clause)

55.1 Clause 4 — Context of the Organization

Intent: Understand context, interested parties, and define AIMS scope.

Minimums: Interested-parties & obligations registers; approved scope; diagrams; trust zones.

Implement: Workshops → obligations register → scope statement → diagrams with versioning and change triggers.

- **Evidence:** Context report; registers; approved scope; diagrams; version history.
- **Acceptance Criteria:** Registers current (annual or on material change); scope artifacts change-controlled.
- **Common Failures:** Over-narrow scope; missing customer/regulatory commitments.
- **Internal QA:** Annual refresh + material-change trigger; sample two use cases.
- **Docs:** AIMS Scope Statement; Interested Parties Register; Obligations Register.

5.2 Clause 5 — Leadership

Intent: Demonstrate commitment; set policy, objectives, roles, resources.

Minimums: AIMS policy & objectives; governance calendar; risk acceptance rules; KPI/KRI dashboards.

Implement: Steering committee; exception/expiry governance; comms plan; resource critical roles.

- **Evidence:** Signed policy; role matrix; minutes; resourcing actions; comms.
- **Acceptance Criteria:** Governance on schedule; decisions & exceptions documented with expiry; resourcing tracked to closure.
- **Common Failures:** Policy without resourcing; unmanaged exceptions.
- **Internal QA:** Quarterly spot-check of three decisions for evidence quality.
- **Docs:** AIMS Policy; Role/Responsibility Matrix; Governance Charter.

5.3 Clause 6 — Planning (Risks, Opportunities, Objectives)

Intent: Address risks/opportunities; set measurable objectives; plan changes.

Minimums: Risk criteria & workflow; treatments; objectives with baselines/targets; change impact analysis.

Implement: Taxonomy (safety, reliability/factuality, privacy/PII, IP/duplication, bias/fairness, transparency/consent, security/abuse, financial/legal/regulatory); appetite; KRIs; tiered change gates.

- **Evidence:** Risk register; OKRs & KRIs; change records.
- **Acceptance Criteria:** All T3 risks owned with treatments/dates; residual risk within appetite; objectives tracked.
- **Common Failures:** Static register; no measures; unmanaged change.
- **Internal QA:** Monthly risk; quarterly objectives.
- **Docs:** Risk Management Procedure; OKRs; Change Impact SOP.

5.4 Clause 7 — Support

Intent: Resources, competence, awareness, communication, document control.

Minimums: Role-based competences; onboarding before access; annual/targeted refresh; doc lifecycle; comms plan.

Implement: Competence matrices; access gating on training; approvals/versioning; internal/external comms.

- **Evidence:** Rosters & assessments; competence verification; doc approvals; comms.
- **Acceptance Criteria:** ≥98% completion; competence verified; docs current & approved.
- **Common Failures:** Shadow docs; outdated SOPs; inadequate role training.
- **Internal QA:** Quarterly doc control audit; semiannual competence sampling.
- **Docs:** Competence & Training Plan; Document Control SOP; Communications Plan.

5.5 Clause 8 — Operation

Data governance: Rights/provenance; classification/retention; PII minimization; dataset versioning & lineage; allowed sources; quality gates.

TEVV (tiered):

- **T1:** Factuality; harmful content; PII leak; basic jailbreak.
- **T2:** + Robustness (adversarial, tool-abuse, injection chaining); bias slices; regression per release.
- **T3:** + Red-team; domain safety; membership-inference/duplication; legal/IP; HIL gate; rollback drills; pre-deployment approval board.

Deployment & guardrails: Filters; grounding/citations; allow/deny; abstain/escalate/ask-human; rate limits; safety metadata; signed artifacts; secrets; isolation; logging & tamper protection.

Human oversight & transparency: Reviewer criteria; escalations; disclosures; decision logs.

Incident handling: Taxonomy; severity SLAs; AARs; effectiveness checks.

Suppliers: Responsibility matrices; eval/telemetry rights; incident/change notice; IP/indemnity; data usage/retention; subcontractor flow-downs.

Release Gate Matrix (by tier)

Tier	Required Evals	Minimum Passing Thresholds (examples)	Sign-off Roles	Rollback Plan
T1	Factuality, harmful, PII, basic jailbreak	Unsupported ≤5%; jailbreak ≤2%; 0 confirmed PII	Model Owner + AIMS Mgr	Basic rollback documented
T2	T1 + robustness, bias, regression	Unsupported ≤2%; jailbreak ≤1%; fairness delta defined	Model Owner + AIMS + Sec/Privacy	Tested rollback in non-prod
T3	T2 + red-team, domain safety, membership-inf., IP	Unsupported ≤2%; jailbreak ≤0.5%; 0 PII; pass domain safety	Approval Board (incl. Legal)	Live rollback drill ≤7 days

Deviation & Compensating Controls: Deviations approved by AIMS Mgr + Exec Sponsor; compensating controls documented with expiry and tracked in an exception ledger.

Evidence: Rights attestations; registries; TEVV plans/reports; red-team results; guardrail configs; dashboards; incidents/AARs; supplier matrices & reviews.

- **Acceptance Criteria:** Thresholds met or deviations with expiry; guardrails default-on; HIL triggers tested; incident SLAs met.
- **Common Failures:** Missing rights proof; no jailbreak/PII/IP tests; untested rollback; unmanaged vendor changes; weak HIL.
- **Internal QA:** Quarterly attack simulation; rollback drill; supplier review with actions.
- **Docs:** Data Governance Standard; TEVV Plan; Deployment & Guardrails SOP; Incident Response (AI); Supplier Responsibility Matrix.

5.6 Clause 9 — Performance Evaluation

KRIs: Safety violations; unsupported claim rate; grounded hallucination; jailbreak success; PII/IP leakage; fairness deltas; MTTA/MTTR; guardrail coverage; change frequency; restore success; logging coverage; provider-notice SLA.

Internal audit: Risk-based annual coverage of Clauses 4–10; impartial auditors; sampling & re-performance; CAPA verification.

Management review: Strategy fit; objectives; KRI trends; audit/nonconformity summaries; incidents; suppliers; resources; changes; opportunities.

Acceptance Criteria: KRIs trended; audits executed; MR actions closed with evidence.

Common Failures: Checklist audits; perfunctory reviews; metrics without thresholds/owners.

5.7 Clause 10 — Improvement

Cycle: Detect → Contain → RCA → Plan → Implement → Verify effectiveness → Close → Update risks/docs/training.

Acceptance Criteria: On-time closures; decreasing recurrence; metrics improve.

Internal QA: Quarterly review of repeat issues & overdue actions.

Docs: Nonconformity & Corrective Action Procedure; Continual Improvement Plan.

5.8 Applied Narratives by Tier (T1 & T3)

T1 Narrative (Internal Productivity Bot): Small prompt change → Change ticket (T1) → Minimal TEVV (factuality + PII + basic jailbreak) meets thresholds → Model Owner + AIMS sign-off → Monitoring confirms no regression in a week → Record in Model Registry.

T3 Narrative (Customer-Facing Agent Launch): New provider + major autonomy jump → Full impact analysis → T3 TEVV incl. red-team, domain safety, membership-inference, IP → Approval Board sign-off → **Rollback drill within 7 days** → Continuous KRIs → MR reviews trend; any deviation requires compensating controls with expiry.

6. AI Authorization Boundary & Workflow Mapping

- **Views:** (1) Context & trust zones; (2) Data lineage; (3) Model/prompt lifecycle; (4) Guardrail stack; (5) Telemetry & logs; (6) Change authorities.
- **Acceptance:** Diagrams exist & controlled; link to logging/TEVV/rollback; access restricted; review date visible.

7. Canonical Documentation Set (AIMS Narratives)

AIMS Manual; Policies & Standards; Procedures & Work Instructions; Records (risks, evaluations, incidents & AARs, audits, reviews, corrective actions, supplier reviews). Maintain an indexed catalog with owners, retention, last-updated.

8. Applicability, Inheritance & Acceptance Criteria Mapping

Track per clause: applicability, implementation summary, ODPs, measurable acceptance criteria, inheritance (if any), evidence link, **status** (C/PC/NC/N/A).

Example — Clause 8 (as above).

9. Organization-Defined Parameters (ODPs)

Domain	T1	T2	T3	Owner	Review Cadence
Jailbreak / Prompt-Injection Success	≤2%	≤1%	≤0.5%	TEVV Lead	Monthly
Unsupported Claim Rate (grounded tasks)	≤5%	≤2%	≤2%	Model Owner	Monthly
PII Leakage	0 confirmed	0 confirmed	0 confirmed	Sec/Privacy Lead	Continuous
Fairness Disparity (pp)	≤5 (unless justified)	≤5	≤5	TEVV Lead	Quarterly
Rollback Drill	N/A	Semiannual	Within 7 days post-release	SRE/Model Ops	Per release/Tier
Incident SLA (High)	Triage ≤1h	≤1h	≤1h	IR Lead	Continuous
Logging Retention	≥12 months	≥12 months	≥12 months	SRE/Model Ops	Annual

Apptega

10. Evidence Register

Staleness rule: Evidence without a **Last Updated** or **Verifier** is treated as stale.

Artifact	Clause(s)	Link/Path	Owner	Format	Last Updated	Verifier	Sample Size	Next Review
AIMS Policy & Scope	4,5	AIMS/Policy/	AIMS Mgr	PDF				
Context & Interested Parties	4	AIMS/Context/	AIMS Mgr	PDF				
Obligations Register	4	AIMS/Obligations/	Legal/Ethics	CSV/PDF				
Architecture & Data Flows	4,8	AIMS/Arch/	Eng Lead	PDF				
Risk Method & Register	6	AIMS/Risk/	Risk Owner	CSV/PDF				
Objectives, KPIs/KRIs	6,9	AIMS/Metrics/	AIMS Mgr	CSV/PDF				
Training Rosters & Assessments	7	AIMS/Training/	HR/L&D	CSV/PDF				
Document Control Records	7	AIMS/Docs/	Quality	CSV/PDF				
Dataset Registry & Rights	8	AIMS/Data/	Data Owner	CSV/PDF				
Model Registry (versions, approvals)	8	AIMS/Models/	Model Owner	CSV/PDF				
TEVV Plans & Reports	8	AIMS/TEVV/	TEVV Lead	PDF/CSV				
Red-Team Results	8,9	AIMS/RedTeam/	Sec/TEVV	PDF				
Guardrail & Policy Configs	8	AIMS/Deploy/	Model Ops	JSON/PDF				
Monitoring Dashboards & Logs	9	AIMS/Monitoring/	SRE/Model Ops	PDF/CSV				
Incident Tickets & AARs	8,10	AIMS/IR/	IR Lead	PDF				
Corrective Action (RCA, effectiveness)	10	AIMS/CA/	AIMS Mgr	CSV/PDF				
Supplier Inventory & Matrices	8,15	AIMS/TPRM/	Procurement	PDF/XLSX				
Supplier Due Diligence & Reviews	8,15	AIMS/TPRM/	Procurement	PDF				
Internal Audit Program & Reports	9,10	AIMS/Audit/	Audit Lead	PDF				
Management Review Minutes & Actions	9	AIMS/MR/	Exec Sponsor	PDF				

11. Continuous Monitoring (AIMS)

KRI Examples & Alerting

KRI	Sample Threshold	Owner	Alert Route	Auto-ticket
Jailbreak success (T3)	≥0.5% weekly	TEVV Lead	Pager/Email	Yes
Unsupported claim rate (critical use case)	≥2% weekly	Model Owner	Slack/Email	Yes
PII leakage alerts	Any confirmed	Sec/Privacy	Pager	Yes
Restore failure (monthly drill)	Any failure	SRE/Model Ops	Pager	Yes
Vendor notice SLA misses	>0 in quarter	Procurement/TPRM	Email	Yes

- **Workflow example:** Jailbreak spike → SIEM rule → auto-ticket → TEVV hotfix → guardrail update → re-eval → closure with metrics & timeline → MR review.
- **Cadence:** Daily log health; weekly bias/RAG/secret checks/prompt diffs; monthly threshold & exception review + restore drill; quarterly red-team, rollback, KRI/MR supplier review.
- **Acceptance:** KRIs have thresholds, owners, trend lines; breaches ticketed and closed within SLA.

11.1 Metric Definitions & Calculation Notes

- **Unsupported Claim Rate (UCR):** (# responses with claims not supported by designated ground truth / # evaluated responses) × 100.
- **Jailbreak Success Rate (JSR):** (# successful policy-bypass outcomes / # adversarial attempts) × 100 (track internal vs external).
- **PII Leakage Confirmed:** Count of incidents where output contained regulated PII verified by Sec/Privacy (near-misses tracked separately).
- **Fairness Disparity (pp):** Max absolute difference of outcome rates across defined slices for the same task.
- **Hallucination (Grounded):** (# incorrect factual assertions contradicted by sources / # evaluated grounded responses) × 100.
- **Guardrail Coverage:** % of live routes/endpoints with enforced filters + logging validated by probes.

12. Internal Audit Program (AIMS)

- **12-month plan (illustrative):** Q1: Clauses 4–6 • Q2: Clause 8 • Q3: Clause 9 • Q4: Clause 10 & supplier deep-dive.
- **Sampling recipe:** ≥10 items **or** ≥10% (whichever higher), stratified by tier and recency.
- **Finding types & closure:** Major (RCA ≤48h; effectiveness ≤30/45/60 days H/M/L) vs Minor (verify by next quarter).
- **Reporting:** CAPA register with effectiveness verification and re-performance.
- **Acceptance:** Plan executed; majors closed on time; recurrence trend down.

13. Management Review (AIMS)

- **Agenda (90 minutes):** Strategy & policy fit; Objectives & KRI trends; Audit & nonconformities; Incident themes; Supplier performance; Resource needs; Risk acceptances & exceptions (expiries); Decisions/action review.
- **Packet:** dashboards, audit summary, CAPA aging, exception ledger with expiries, supplier notices, resourcing asks.
- **Cadence:** Semiannual minimum; plus after material change or major incident.
- **Acceptance:** Minutes list decisions/actions; 100% actions closed or extended with justification by next review.

14. Corrective Action Workflow

- **Timelines:** High (triage ≤1h; containment ≤24h; RCA ≤48h; effectiveness ≤30d) • Medium (≤4h; ≤3d; ≤45d) • Low (≤1 business day; ≤60d).
- **Effectiveness proofs:** e.g., Hallucination **4.8% → 1.9%** over two releases; Membership-inference **1.2% → 0.2%** after data minimization & guardrails.
- **Acceptance:** All Highs have RCA + verified effectiveness; extensions have expiry & approval.

15. Third-Party & Supply Chain (AI Services & Models)

- **Contract Minimums:** evaluation & telemetry rights; incident/change notices; update transparency; IP/indemnity; data usage/retention/deletion; subcontractor flow-downs; right-to-audit; logging retention parity.
- **Onboarding/Offboarding:** Responsibility matrix; due diligence; security & privacy questionnaires; telemetry integration; data-use constraints; then access revocation, deletion certs, final telemetry export, de-scoping in Evidence Register.
- **Acceptance:** Matrices current; due diligence & reviews complete; notices acted upon within SLA; access recerts & revocations evidenced.

16. Competence & Awareness (Role-Based) — with Rubrics

Role	Must Demonstrate	Assessment
TEVV Lead	Design safety/robustness/bias/privacy suites; interpret metrics; mitigations	Scenario exam + red-team lab
Model Owner	Configure guardrails; read logs; run regressions; approve gates	Hands-on practical + quiz
SRE/Model Ops	Secret mgmt; rollback drills; logging health; restores	Live runbook re-perf
Reviewers/HIL	Apply decision criteria; escalate; document	Case review
Sec/Privacy	Classify incidents; validate PII/IP; approve notices	Incident tabletop
Legal/IP & Ethics	Rights/duplication checks; contract clauses; exception review	Contract review
AIMS Manager	Maintain ODPs/evidence; exception ledger; MR readiness	Portfolio walkthrough

Thresholds: ≥98% completion; critical roles ≥85% pass; downward trend in repeat findings.

17. Change Management for AI Lifecycle Impact

Triggers: Changes to datasets, prompts/policies, models, guardrails, tools/functions, RAG sources, dependencies, or providers.

Risk-Tiered Gate Matrix

Trigger × Tier	Impact Analysis	TEVV Delta	Approvals	Rollback Drill	Comms/Training
T1: Prompt/policy tweak	✓	Minimal	AIMS Mgr	N/A	As needed
T2: Model/pipeline update	✓	Moderate	AIMS Mgr + Model Owner	≤30 days	Yes
T3: New model/provider	✓	Full	Approval Board incl. Legal	≤7 days	Yes (targeted)

Acceptance: 100% changes show approvals, TEVV results, and (for T3) a recent rollback drill.

18. Evidence Sampling Plans (Internal QA)

Rules: Scope & diagrams (100% T3; 25% T1/T2); Risk register (top-5 high risks re-scored; treatments verified); TEVV/guardrails (one end-to-end eval quarterly; jailbreak/PII/IP on top two T3 use cases); Incidents (100% High; 50% Medium); Suppliers (two matrices + one notice per quarter; one offboarding annually); Docs & training (five SOPs; 10% of critical-role users).

Pass/Fail & Escalation

Check	Pass	Fail	Escalation
Evidence present & current	Dated & verified	Missing/stale	Raise CAPA
Threshold met	Meets ODP	Below ODP w/o exception	Exception with expiry
Actions on time	Within SLA	Overdue	Exec Sponsor review

19. Common Pitfalls & Anti-Patterns

Paper-only AIMS; missing data rights; weak TEVV; no rollback; unclear supplier duties; logging without review; no exception expiry; audits without effectiveness checks; unmanaged prompt/policy drift; undefined fairness slices.

20. Cloud & Hosting Scenarios — Shared-Responsibility Mapping

Area	SaaS LLM API	Self-Hosted OSS Model	Vendor-Hosted Fine-Tune	On-Prem RAG over Enterprise Data
Data Rights & Provenance	Vendor content policy + your input rights attestation	Full chain of custody owned by you	Split: your data rights; vendor base model license	Your source rights; infra access controls
TEVV Depth	Moderate (vendor attest + your evals)	High (full suite)	High on your training data; vendor baseline	High on grounding, leakage, hallucinations
Guardrails	API filters + app policies	In-app filters + runtime policies	Split (vendor + app)	App + retrieval filters
Logging & Telemetry	App logs; vendor transparency reports	Full stack in your SIEM	Split visibility; require telemetry rights	App + retriever + vector DB logs
Incident Handling	Contractual notice; ticket integration	Internal IR	Dual-party IR	Internal IR; data owner involvement
Monitoring Owner	App team	You	Split	App + Data teams
Primary Logs/ Signals	API logs, rate limiting, vendor reports	Model/runtime logs, TEVV outputs	API + fine-tune job logs	Retriever queries, vector DB, grounding diffs

Acceptance: Responsibility matrix per scenario with clear monitoring ownership and evidence paths.

21. Evidence & Documentation Checklist

- [ ] Scope & boundary diagrams current and versioned
- [ ] Obligations register maintained and reviewed
- [ ] Risk register active with owners & due dates
- [ ] TEVV plan and latest reports per tier
- [ ] Guardrail configurations archived with change history
- [ ] Supplier matrices, due diligence, and notices on file
- [ ] Monitoring dashboards with KRI trends
- [ ] Incident tickets with AARs and corrective actions
- [ ] Internal audit plan, reports, and verified closures
- [ ] Management review minutes with action tracking

22. Training & Awareness Program (AIMS)

- **Structure:** Role-based onboarding; annual refresh; targeted post-incident refresh; micro-learning on jailbreaks, PII leakage, bias, prompt security; simulated drills.
- **Measurement:** Completion %, assessment scores, time-to-train, repeat finding rate, phishing/jailbreak test performance.
- **Acceptance:** ≥98% completion; critical roles ≥85% assessment; downward trend in repeat findings.

23. FAQs (Auditor & Executive)

- **What is "good enough" for Tier 1?** Baseline TEVV on scoped tasks, basic jailbreak probes, PII checks, guardrails enabled with monitoring.
- **How to prove data rights for training & RAG?** Rights attestations per dataset/source; link to contracts/licenses/consent records; store in Dataset Registry.
- **How often to red-team?** Quarterly for T3; semiannually for T2; on material change.

24. Operating Model & Continuous Improvement

- **Cadence:** Weekly KRI review; monthly objective & exception review; quarterly red-team/rollback; semiannual MR; annual internal audit cycle.
- **Decision loop:** Monitoring → Incidents → Audits → Reviews → Update thresholds/controls/training → Track actions to closure with evidence.

25. Quick Reference Summary (Policy Area / Type / Owner / Tools / Examples)

Policy Area	Type	Owner	Primary Tools/ Systems	Examples
Governance & Policy	Administrative	Exec Sponsor / AIMS Mgr	GRC, Docs	AIMS Policy; Governance Charter; Risk Appetite
Risk Management	Admin/Technical	Risk Owner	GRC, Risk Reg	Risk Method; Register; KRI set
Data Governance	Admin/Technical	Data Owner	DLP, DG	Rights attestations; lineage; retention
TEVV	Technical	TEVV Lead	Eval harness, Red-team	Safety/factuality/bias/privacy evals
Deployment & Guardrails	Technical	Model Owner / SRE	CI/CD, Sec tools	Filters; allow/deny; secrets
Incident Mgmt	Admin/Tech	IR Lead	Ticketing/SIEM	Taxonomy; IR plan; AARs; CAPA
Supplier Mgmt	Admin/Contract	Procurement/TPRM	TPRM	Responsibility matrices; notices
Training & Competence	Administrative	HR/L&D	LMS	Curricula; assessments
Monitoring & Audit	Admin/Tech	SRE / Audit Lead	SIEM/GRC	Dashboards; internal audits
Improvement	Administrative	AIMS Mgr	CAPA tracker	RCA; effectiveness checks

26. Self-Assessment & Leadership Attestation

Self-Assessment Grid: Track C/PC/NC/N/A, owner, evidence link, action ID for each clause and practice.

Leadership Attestation (paste onto letterhead):
I attest that:

1. The AIMS scope is complete and accurate for the period stated.
2. Evidence exists for all items marked Compliant.
3. PC/NC items have assigned owners and due dates.
4. Risk acceptances are documented with expiry and compensating controls.

Signature

Name/title

Date

Re-attest after each MR or material change.

27. References & Resources

ISO/IEC 42001:2023
https://www.iso.org/standard/81230.html

ISO/IEC 23894:2023
https://www.iso.org/standard/77304.html

ISO/IEC 27001:2022
https://www.iso.org/standard/82875.html

NIST AI RMF 1.0
https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

OWASP Top 10 LLM
https://owasp.org/www-project-top-10-for-large-language-model-applications/

C2PA
https://c2pa.org

28. Glossary & Acronyms

- **AAR** (After-Action Review)
- **IMS** (AI Management System)
- **CAPA** (Corrective and Preventive Action)
- **HIL** (Human-in-the-Loop)
- **JSR** (Jailbreak Success Rate)
- **KPI/KRI** (Key Performance/ Risk Indicator)
- **ODP** (Organization-Defined Parameter)
- **PII** (Personally Identifiable Information)
- **RAG** (Retrieval-Augmented Generation)
- **RCA** (Root Cause Analysis)
- **TEVV** (Testing, Evaluation, Validation, Verification)

29. Templates (Ready-to-Fill)

29.1 Exception & Compensating Controls Ledger

ID	Use Case / System	Deviation	Compensating Control	Approver(s)	Effective From	Expiry	Owner	Evidence Link	Status

29.2 Risk Acceptance Form

Field	Entry
Risk ID / Title	
Description & Context	
Likelihood × Impact (pre-treat)	
Treatment Options Considered	
Residual Risk & Rationale	
Expiry / Review Date	
Approver (Role/Name)	
Evidence / Attachments	

29.3 Supplier Responsibility Matrix (AI Services & Models)

Area	You	Supplier	Shared	Evidence
Data rights & provenance				
Evaluations & telemetry rights				
Incident notification & SLA				
Model updates/transparency				
Data usage/retention/deletion				
Subcontractor flow-downs				
Right-to-audit				
Logging/retention parity				

Apptega

Annex A — Control Mapping (Summary + Status)

Theme	Where Implemented	Primary Evidence	Status (C/PC/NC/N/A)
Governance & accountability for AI	Clause 5; Sec. 5–6, 13	Policy; minutes; role matrix	
AI risk mgmt & objective setting	Clause 6; Sec. 6, 9	Risk method/register; KRIs	
Context, scope & stakeholders	Clause 4; Sec. 2, 6	Scope statement; obligations; diagrams	
Data governance & rights	Clause 8; Sec. 7–8, 15	Rights attestations; lineage	
Model lifecycle (design→TEVV→deploy)	Clause 8; Sec. 5, 8, 17	TEVV reports; registry; rollback	
Safety/robustness/abuse resistance	Clause 8; Sec. 8–11	Red-team; guardrails	
Privacy & confidentiality	Cl. 6–8; Sec. 7–9	PII handling; alerts; AARs	
Bias/fairness & human oversight	Clause 8; Sec. 7–9	Bias tests; HIL	
Transparency & records	Cl. 4–7,9; Sec. 5,7,21–26	Doc control; audit reports	
Incident & corrective action	Cl. 8,10; Sec. 11, 14	Tickets; RCA/CAPA	
Supplier & outsourced processes	Clause 8; Sec. 15	Matrices; contracts; reviews	
Monitoring, audit & review	Clause 9; Sec. 11–13	Dashboards; audits; MR	
Continual improvement	Clause 10; Sec. 14, 24	Improvement register	
Change management	Cl. 6,8; Sec. 17	Change tickets; TEVV deltas; rollback	
Provenance & output integrity	Clause 8; Sec. 5, 8	Provenance logs; labeling policy	

Annex B — Sample Auditor Evidence Pack (What to Hand Over)

Topic	Primary Artifacts	Suggested Sample
Scope & Boundary	Scope statement; diagrams; version history	All T3 + 2× T2
Risk Mgmt	Risk method; top-5 risks; treatments	10 items or 10%
TEVV	Latest TEVV per critical use case; thresholds	1 full suite rerun
Guardrails	Config exports; allow/deny lists	2 endpoints
Incidents	Tickets; AARs; effectiveness	100% High; 50% Med
Monitoring	KRI dashboards; alert proofs	Last quarter
Suppliers	Responsibility matrix; notices	2 vendors + 1 offboard
Audit & MR	Reports; minutes; action logs	Most recent
CAPA	RCA + effectiveness	All open + last 3 closed

Annex C — Linkage & Inheritance Matrix (Expanded + Minimum Linkage Checklist)

Minimum Linkage Checklist (tick before declaring inheritance):

- [ ] Risk appetite & criteria documented
- [ ] TEVV thresholds & dashboards defined and trended
- [ ] Supplier responsibility matrices current with AI evaluation & telemetry rights
- [ ] Incident taxonomy includes AI
- [ ] Document control applies to AI artifacts

(Linkages to ISO/IEC 23894, ISO/IEC 27001/27002, NIST AI RMF, ISO 31000, and Transparency/Provenance retained as described previously.)

Annex D — Non-Normative Regulatory Touchpoints (High-Level)

Informational only; not a claim of compliance.

- **EU AI Act:** Risk classification aligns with Tiering (T1–T3); evidence requirements supported by Sections 8–11; supplier transparency connected in Section 15.
- **Sectoral privacy regimes:** PII handling, logging, incident notice (Sections 7–8, 11, 15).
- **Content provenance & labeling:** C2PA linkage via Section 8, Annex C.

Always consult counsel; keep this annex updated as regulations evolve.

Apptega

Apptega Product Features

16+ Security Frameworks

One-Click Reporting

Automated Alerts & Notifications

API & Application Connectors

Automated Framework Crosswalking

Real-Time Compliance Scoring

Restricted Auditor View

Single Sign-On Connectivity

Policy & Plan Templates

Automated Risk Assessments

Document Repository for Artifacts

Multi-Tenant Environment

Apptega

About Apptega

A perennial G2 leader across various cybersecurity categories, Apptega is the cybersecurity and compliance platform purpose-built for security providers who are growing lucrative security practices, creating stickier customer relationships, and winning more business from competitors.

With Apptega's end-to-end platform, IT providers can:

1. Build continuous compliance offerings that set them apart, retain customers, and grow their ARR.
2. Build and manage world-class cybersecurity and compliance programs for their clients.
3. Increase the capacity and efficiency of their existing team so they can service 2–3× more customers.
4. Assess, measure, and manage their own security programs against a standard set of industry-leading frameworks.

To learn more, visit apptega.com

Visit apptega.com