VISIBILITY

Knowing What AI You Actually Have

The 15 Guardrails of AI — Pillar 1 of 5

A Vertex11 Whitepaper

2 guardrails · 15 controls · 10 key

G1 AI Use Case & Asset Inventory (4 controls, 3 key)

G10 Logging, Auditability & Explainability (11 controls, 7 key)

Validated against NIST AI RMF 1.0, the EU AI Act, ETSI TS 104 223, and ISO/IEC 42001:2023

Executive summary

Visibility is the foundation of AI governance, and the pillar most organisations underestimate. Before a single control can be applied, leadership must be able to answer a deceptively simple question: what AI is running in our organisation, what does it touch, and who is accountable for it? Most cannot. AI has entered the enterprise faster than any technology before it — through sanctioned projects, embedded vendor features, and individual employees experimenting with public tools — and it has done so largely off the books.

This paper makes the case for treating AI discovery, inventory, and auditability as first-class capabilities, details the specific risks that accumulate in their absence, and sets out what is required across people, process, and technology to build and sustain them.

The Visibility pillar contains two guardrails. **G1 — AI Use Case & Asset Inventory** is the keystone on which every other guardrail depends: if you don't know a system exists, no control applies to it. **G10 — Logging, Auditability & Explainability** closes the loop: once you know what you have, you need a tamper-evident record of what it does, the ability to explain its decisions, and monitoring that catches drift before it becomes an incident.

Together they carry 15 control statements — 10 of them key — validated against four international standards: NIST AI RMF 1.0, the EU AI Act (as amended by the 2026 Digital Omnibus), ETSI TS 104 223, and ISO/IEC 42001:2023. Every control traces to at least one standard provision.

Why visibility comes first

Every other guardrail in the framework assumes you know the system exists. You cannot classify the data flowing into a model you haven't catalogued, scope the permissions of an agent you don't know is running, monitor an unrecorded workflow, or shut down a tool you can't locate. Visibility is not one control among many; it is the precondition for all of them.

It is also where governance most often quietly fails. An organisation can invest heavily in monitoring, testing, and incident response and still be exposed, simply because a third of its AI footprint never made it onto any list. Coverage gaps, not control weaknesses, are the most common root cause of preventable AI incidents.

Visibility delivers three things at once: a defensible answer to oversight questions from boards, auditors, and regulators; an accurate map for prioritising where to spend limited governance effort; and the early detection of risk before it becomes an incident.

The standards validation reinforced this positioning. NIST's GOVERN function assumes a complete system inventory before any risk management activity begins. The EU AI Act's registration and documentation obligations cannot be met without one. ETSI TS 104 223 requires baseline technical records for every AI system. And ISO/IEC 42001 expects the organisation to identify all AI activities and document their characteristics. Four standards, one message: you cannot govern what you cannot see.

The risks that manifest without it

Shadow AI

Employees adopt public AI tools to get work done, often pasting customer records, source code, or confidential strategy into systems with no data protection guarantees. Each instance is an unmonitored data egress point. Without an inventory, the organisation has no idea it is happening until it appears in a breach report.

Embedded and supply-chain AI

Vendors increasingly add AI features to existing products without fanfare. A CRM, a support tool, or an HR platform can begin processing your data through a model you never evaluated. These additions rarely trigger a procurement review, so the risk enters invisibly through tools you already trust.

Orphaned and unowned systems

AI built for a pilot or a single team frequently outlives its sponsor. Without recorded ownership, no one is accountable for its security, accuracy, or decommissioning, and it continues operating long after anyone is paying attention to it.

Blind risk prioritisation

Governance budgets are finite. Without a complete inventory and risk rating, organisations spread effort evenly or, worse, secure the visible systems while the highest-risk ones go untouched.

Inability to respond

When a regulator asks how many high-risk AI systems you operate, when an auditor requests evidence, or when an incident requires you to identify every affected system, the absence of an inventory turns a manageable request into a crisis.

Invisible decisions without explanation

AI systems increasingly make or influence consequential decisions — credit scoring, recruitment filtering, clinical prioritisation, insurance pricing. Without logging and explainability, affected individuals cannot understand the basis of decisions that affect them, organisations cannot demonstrate compliance, and oversight bodies cannot audit. Even outside a regulatory mandate, an unexplainable decision is an unaccountable one.

Silent drift

A model that was accurate at deployment degrades as the world changes around it. Without monitoring the model can consumes a continuous log stream, where drift goes undetected — accuracy erodes, bias shifts, and the organisation continues to trust outputs that no longer deserve it. By the time anyone notices, the damage is done and often irreversible. In one instance the model may drift, in others the requirements may drift creating another set of mismatched expectations and irreversible impact.

Guardrail 1: AI Use Case & Asset Inventory

4 controls · 3 key | Pillar: Visibility | Applicability: Universal

The inventory is the single most consequential capability in the framework. Not because it is technically complex — it is not — but because without it, nothing else works. Every guardrail from data classification to incident response assumes you know the system exists. An unregistered AI system is a system with no data controls, no access policy, no monitoring, no owner, and no recovery plan.

Guardrail 1's four controls establish a continuously maintained register of every AI system, model, agent, and third-party provider in the organisation. Each entry carries a named accountable owner — because ownership is not optional metadata; it is the mechanism through which accountability flows to every downstream control.

What the controls cover

The controls address four distinct problems. First, the register itself: ensuring one exists and that it captures every AI asset, not just the ones that went through formal procurement. Second, the depth of each entry: purpose, data categories, risk classification, deployment context, and — following ISO/IEC 42001 validation — the dependent tooling, compute infrastructure, and competencies involved across the system's lifecycle. This transforms the inventory from an asset list into a dependency and capability map.

Third, the decay problem. A point-in-time inventory is immediately out of date. The controls require both a defined refresh cadence and an active discovery process to surface AI that bypassed intake — shadow deployments, embedded vendor features, and experimentation that scaled without governance oversight.

Fourth, the gate. No AI system reaches production without an inventory entry and assigned owner. This is the control that prevents the next generation of shadow AI from accumulating.

Guardrail 10: Logging, Auditability & Explainability

11 controls · 7 key | Pillar: Visibility | Applicability: Conditional & Universal

Guardrail 10 is the deepest guardrail in the Visibility pillar — 11 controls, 7 of them key. It exists because knowing what AI you have is necessary but not sufficient. You also need to know what it is doing, be able to prove what it did, explain why it made a particular decision, and detect when it starts behaving differently from what was intended.

The guardrail covers four related capabilities that, together, turn a static inventory into a live governance system.

What the controls cover

Logging and attribution. Every consequential action and decision must be recorded in enough detail to reconstruct what happened and trace it to a responsible identity. The logs themselves must be tamper-evident and retained per policy — because a log that can be altered after the fact provides no assurance at all.

Explainability and transparency. The reasoning behind consequential decisions must be surfaceable and communicable to affected stakeholders in terms they can understand. This is not a requirement for full model interpretability — that would rule out most modern architectures — but a design-time commitment to making decisions intelligible. Separately, the controls require disclosure of where AI is in use, where it should not be relied upon, and what its prohibited use cases are.

Monitoring. Logging without monitoring is an archive. The controls require that the log stream is consumed in near real-time to detect deviation, anomaly, and policy breach as they happen — not discovered weeks later in a periodic review.

The feedback loop. Findings from monitoring, audit, explainability reviews, and stakeholder feedback must be routed back to the inventory and risk classification at Guardrail 1. A channel must exist for affected parties to query or contest decisions. This closes the governance loop: without it, the framework assesses once and assumes indefinitely.

Documentation at the right depth. The controls establish a tiered documentation requirement — from a universal baseline technical record for every AI system, through deployer instructions covering capabilities, limitations, and human-oversight measures, to the full technical documentation depth required by the EU AI Act's Annex IV for high-risk systems. The depth is proportionate to risk; the obligation to document is universal.

Building the capability: people, process, technology

People

Role	Responsibility
AI governance lead	Owns the inventory and logging framework end to end; accountable for completeness, currency, and auditability.
Business-unit AI owners	Register and maintain entries for their area; the front line of accuracy. Responsible for ensuring logging is active on their systems.
Procurement	Flags AI capabilities in new and renewing vendor contracts before they go live. Validates that vendor systems meet logging and documentation requirements.
Security & data office	Supply discovery signals, validate risk ratings, and ensure log integrity and retention meet policy.
Executive sponsor	Mandates participation so the inventory and logging are not optional. Signs off on the AI policy (per ISO 42001).

The decisive factor is clear accountability. An inventory without a named owner and a RACI that reaches into every business unit will decay within a quarter. The same applies to logging — if no one owns the log architecture, log quality degrades until it is useless at exactly the moment you need it most.

Process

Intake and onboarding: every new AI use case — built or bought — is registered before deployment, with registration embedded into existing project and procurement workflows rather than bolted on. Logging requirements are defined at intake, not retrofitted.

Risk-tiering rubric: a consistent method for classifying each use case by impact, data sensitivity, autonomy, and regulatory exposure, so governance effort — including documentation depth — can be proportionate.

Periodic attestation: owners confirm their entries are accurate on a set cadence; stale entries are escalated.

Discovery cadence: scheduled sweeps to surface shadow and embedded AI that bypassed intake.

Feedback loop: monitoring findings, audit results, and stakeholder feedback are routed back to the inventory and risk classification on a defined cadence.

Decommissioning: a defined path to retire systems, dispose of data securely, and remove them from active scope.

Technology

A system of record for the inventory — a dedicated AI governance platform, a GRC module, or an extended CMDB — rather than a spreadsheet that no one trusts.

SaaS and network discovery tools to detect AI usage from traffic, expense, and identity signals.

Centralised, tamper-evident logging — SIEM integration, immutable log stores, or integrity-anchored hashing — that ensures logs cannot be altered after the fact.

Monitoring and alerting that consumes the log stream in near real-time, with defined thresholds for deviation, anomaly, and policy breach.

Explainability tooling appropriate to the model types in use, embedded in the development and deployment process rather than retrofitted at audit time.

A defined metadata schema so every entry captures the same fields and the inventory is queryable and reportable.

How the standards shaped this pillar

The Visibility pillar existed before any of the four standards were applied. What the validation did was pressure-test the controls against the specific expectations of each standard — strengthening the pillar from standard to standard.

NIST AI RMF 1.0 reinforced the feedback loop. The framework already had monitoring; NIST made the return path to the inventory explicit and mandatory.

The EU AI Act brings documentation depth and requires deployers to receive clear instructions for use. These are not aspirational; they are conditions for market access.

ETSI TS 104 223 closed operational gaps by embedding baseline technical records in the inventory. Practical capabilities many organizations are missing in production today.

ISO/IEC 42001 widened the inventory capturing compute infrastructure, dependent tooling, people / competencies involved and it requires plain-language system documentation for all users, not just deployers of regulated systems.

The net effect: the Visibility pillar now covers the inventory, logging, explainability, monitoring, and documentation expectations of all four standards. Organisations implementing it can demonstrate compliance across NIST, EU, ETSI, and ISO requirements from a single set of controls.

A maturity path

Most organisations begin with a manual, point-in-time survey — useful to start, but immediately out of date. The progression runs through four levels:

Level 1 — Manual survey. A spreadsheet-based inventory with no logging infrastructure. Governance relies on periodic manual reviews. This is where most organisations are today.

Level 2 — Centrally owned register with intake gating. A dedicated system of record with mandatory registration before deployment. Basic logging is in place for high-risk systems. Documentation requirements are defined but compliance is inconsistent.

Level 3 — Integrated and partially automated. Discovery tools detect shadow AI. Logging is centralised with tamper-evidence. Monitoring alerts on anomalies. Explainability is available for consequential decisions. The feedback loop to the inventory operates on a defined cadence.

Level 4 — Continuously reconciled. New AI is detected and registered with minimal human effort. Monitoring is real-time. Explainability is embedded in the development process. Documentation is

generated from the system of record. The feedback loop is continuous and automated where appropriate.

The aim is not perfection on day one, but a system that is trusted, current, and complete enough to drive the controls in every pillar that follows.

Connecting to the rest of the framework

Visibility feeds everything downstream. The risk rating assigned at Guardrail 1 determines how much protection, oversight, and assurance each system receives. The logging and monitoring at Guardrail 10 provides the evidence base that the Assurance pillar's testing and attestation controls depend on. The inventory is the map that Resilience & Recovery depends on: when an incident occurs, it tells you which systems are affected, who owns them, and what their dependencies are.

The feedback loop at Guardrail 10 creates a two-way connection: downstream findings — from testing, incidents, audit, and stakeholder feedback — flow back to update risk classifications and inventory entries. Governance is not linear; it is a cycle, and Visibility is both the starting point and the return address.

About this series

This is Pillar 1 of The 15 Guardrails of AI, published by Vertex11. The series covers the full framework across five whitepapers — Visibility, Protect, Oversight, Assurance, and Resilience & Recovery — each aligned to the guardrails, controls, and standards mapping in the V11 AI Governance Framework.

To assess your current AI visibility — including the shadow and embedded AI you may not know about — get in touch.