E-guide

AI data exposure playbook: stop leaks across LLMs, SaaS & pipelines

Visibility and control for AI data across LLMs, SaaS, and pipelines, without slowing teams

relyanceai

Table of Contents

1 Introduction

AI is now embedded everywhere, LLMs in products, "assistants" inside SaaS, RAG pipelines, and logs that quietly copy sensitive data into places it doesn't belong. Traditional privacy and security controls weren't built for **AI data exposure**: prompts and outputs move fast, vector databases persist context, and third-party features change without notice.

This playbook gives you a **practical, end-to-end method** to regain control. You'll learn how to **discover** real AI usage (including shadow AI), **map** sensitive flows into and out of models, SaaS, and pipelines, and **enforce** AI-specific guardrails, so leaks are prevented, obligations are met, and evidence is always audit-ready.

What's inside:

- A step-by-step approach to inventory AI systems, classify sensitive data, and trace flows across code > cloud > model I/O > logs/vector DBs > downstream SaaS.
- Guardrails you can operationalize fast routing/minimization, redaction at model boundaries, least-privilege for AI data stores, and drift detection tied to contracts and policy.
- A 30/60/90 plan, templates (AI System Register, policy examples), and KPIs to prove progress.

Use this guide to **stop LLM data leaks, reduce risk, and show compliance, without slowing teams down.**

Who this guide is for

Security & Data Leaders (CISO/DSPM owners):
Need real-time visibility and policy enforcement for AI-era data flows.

Privacy & Legal (DPO/Privacy Counsel):
Need Universal ROPAs, DPIAs, and alignment between contractual obligations and real processing.

Engineering & Platform/ML:
Need "privacy-as-code" and dev-friendly controls without slowing delivery.

2

Part I

Discovery: Build a complete inventory of AI usage

Goal

Identify every place AI interacts with sensitive data, your own models, embedded AI in SaaS, and pipelines (training, inference, logs, analytics, vector DB/RAG).

Actions

Start by enumerating AI systems and touchpoints across code, APIs, SaaS, infrastructure, and model endpoints, including any "shadow AI" introduced by vendors or teams. Detect sensitive data in motion prompts, outputs, logs, embeddings, and analytics exports and record who and what accesses it. Catalog the vendors and regions that receive AI-related data, along with legal bases and purposes, so Universal ROPAs can reflect operational reality rather than static assumptions. Finally, link what you've discovered to obligations by aligning processing with contractual and regulatory requirements.

Deliverables

- AI System Register (first/third-party)
- Sensitive Data Catalog (by system, purpose/legal basis)
- Vendor & Residency Matrix + initial obligation links (for DPA/SCC alignment)

3

Part II

Map: Trace sensitive flows into/out of LLMs & AI-enabled SaaS

Goal

Turn discovery into a **live data-journey map** that shows where sensitive attributes enter, transform, exit, and persist.

Actions

Build a data-journey graph from source code > services/APIs > model I/O > logs/analytics > SaaS vendors > human/app consumers, with context on how data moves, who accesses it, and why. Overlay obligations on each edge DPAs, SCCs, purpose limitations, and internal policies and highlight mismatches.

AI-specific hotspots:

- Training ingestion (raw dumps > features > training corpora)
- Prompt/response streams (PII/secrets in prompts; PII echo in outputs)
- RAG/vector DB (sensitive snippets in embeddings; broad similarity search)
- AI add-ons in SaaS (over-sharing fields to “assistants”)
- Observability/analytics (prompt/response logs leaving safe boundaries)
- These are precisely the areas where Relyance stresses data-in-use/in-motion visibility.

Deliverables

- Live AI Data Journey Map with obligations overlay
- Gap list: flows that contradict contract terms or stated purposes

4

Part III

Enforce: Guardrails that prevent, detect, and prove

Goal

Close the loop with controls that act on the map, automatically.

- **Prevent (before the leak):** Route and minimize data so only permitted types reach approved regions/vendors, and minimize fields sent to AI assistants in SaaS. Apply redaction/masking at ingress and egress to filter regulated attributes at prompt and log boundaries. Enforce least privilege on vector stores and training artifacts. Embed developer-native checks in CI/CD via source-code analysis (privacy as code).
- **Detect (when reality drifts):** Alert on contract-to-reality mismatches. Monitor data-in-use/in-motion and AI behavior to spot violations or unusual access patterns. Trigger DPIAs automatically when new high-risk processing emerges.
- **Respond & Prove (close the loop):** Orchestrate remediation, pause a flow, rotate keys, update policies integrated with existing tools and workflows. Maintain Universal ROPAs and assessment records that reflect real processing, not stale forms.

Deliverables

- Guardrail policy library (prevention & detection rules)
- Runbooks (who/what/when) connected to existing systems
- Audit pack (current ROPAs + DPIAs + contract-flow diffs)

5 30/60/90-day rollout plan

Days 0–30 (Foundations)

- Connect repos, SaaS, data stores, and AI endpoints; inventory all AI usage and vector/RAG components.
- Stand up live data-journey maps and seed Universal ROPAs from observed processing.

Days 31–60 (Controls)

- Enforce first guardrails: field minimization/routing, redaction at prompt/log edges, least-privilege for vector DBs; wire alerts to SIEM/dev workflows.

Days 61–90 (Operationalize & prove)

- Expand to long-tail SaaS and secondary pipelines; enable automated DPIA triggers on flow changes; finalize audit pack.

KPIs

- % AI systems discovered & mapped: Current: 0% → 30-day target: ≥85% > 90-day: ≥95%
- % flows with obligation links: Current: 0% → 60-day target: ≥90% > 90-day: ≥95%
- Contract–reality drifts detected; MTTR: Target: trend ↓ ; MTTR ≤ 48h (≤24h for high-risk)
- ROPA/DPIA freshness: Target: ≤30 days avg; ≤7 days for high-risk processing

6 Templates & checklists

AI System Register (sample fields)

System • Owner • Type (Train/Infer/Assist) • Data Categories • Purpose/Legal Basis • Vendors/Regions • Storage/Logs • Controls Applied • ROPA Link • Last Reviewed

Per-system checklist:

- ☑ All fields populated
- ☑ Purpose & legal basis documented
- ☑ Vendor/region validated against contracts
- ☑ Controls applied (redaction, minimization, access)
- ☑ ROPA link added and accessible
- ☑ Last reviewed ≤ 90 days ago

Guardrail policy ideas

- LLM-Prompt PII Filter: Block or mask regulated categories unless purpose/legal basis/region/vendor all pass.
- Vector-DB Access Scope: Service-account only; approvals for cross-project reads; disallow ad-hoc analyst export.
- Contract Mismatch Alert: Fire when a data category flows to a vendor where the DPA disallows or omits it.

7

Why Relyance AI aligns to this playbook

- **AI-native visibility from code > cloud > AI models** (full data journeys, real-time).
- **Universal ROPAs** built from actual processing across jurisdictions (no guesswork forms).
- **Unified obligation/contract analysis** that links legal commitments to live data flows.
- **Unified trust & governance system** with detection + **orchestrated remediation**; works with your existing stack.
- **AI-privacy automation** to discover AI systems and maintain audit trails for AI processing.