ThreatLens

WHITEPAPER

The Intelligence Layer for Enterprises

A governance model for adopting AI without losing control of your data.

Executive summary

Enterprises face a dilemma. AI drives real productivity, but every prompt can carry regulated or confidential data into systems the business does not control. Banning AI drives it underground; allowing it freely exposes the company. ThreatLens offers a third path — **govern it**. This paper describes the governance model: a single layer in front of every AI request that classifies the content, applies the policy security defines, protects sensitive data, routes to trusted destinations, and records an immutable decision. The result is AI adoption everywhere, with provable control.

1. The problem: shadow AI and the data-loss surface

Employees adopt public AI tools faster than security can review them. Each prompt may contain personal data, payment information, source code, secrets, or board strategy — leaving the building with no record and no control. When auditors or regulators ask what data went where, there is no answer. Model-native controls, where they exist, are inconsistent from vendor to vendor and are never under the enterprise's own authority.

2. Why existing approaches fall short

Block it	Network and proxy bans drive usage underground and forfeit the productivity gains.
Allow it	An acceptable-use policy with no technical enforcement is hope, not control.
Legacy DLP	Built for email, endpoints, and files — blind to conversational AI intent and context, and noisy.
Vendor controls	Fragmented per provider; the enterprise owns neither the policy nor the record.

3. The ThreatLens approach: a governance choke point

Every request — from the workspace or through the gateway API — passes through one place that does six things, in order, before a model is ever touched:

3.1 Classify the content — Each request is classified into one of ten data classes using deterministic detectors (secrets, payment cards, identifiers, injection patterns) and a semantic classifier for business-sensitive content. Classification runs on a fixed, safe classifier — never on the chat destination.

3.2 Apply the policy matrix — For every data class, security sets a minimum trust tier, an action, and an internet-access mode. One grid expresses the whole policy, and it is owned by security — not by individual users.

3.3 Route by trust tier — Destinations are ranked by trust: public frontier, enterprise-managed (your own keys), customer-managed, and private / local. Confidential classes route only to a destination that meets the required tier — otherwise the request is blocked.

3.4 Protect with DLP — Sensitive spans are redacted inbound and outbound. Raw secrets and prompt-injection are always blocked, regardless of how the matrix is configured.

3.5 Ground answers safely — Governed retrieval answers from your Microsoft 365 content, access-trimmed to exactly what the requesting user is permitted to see, and fail-closed by default.

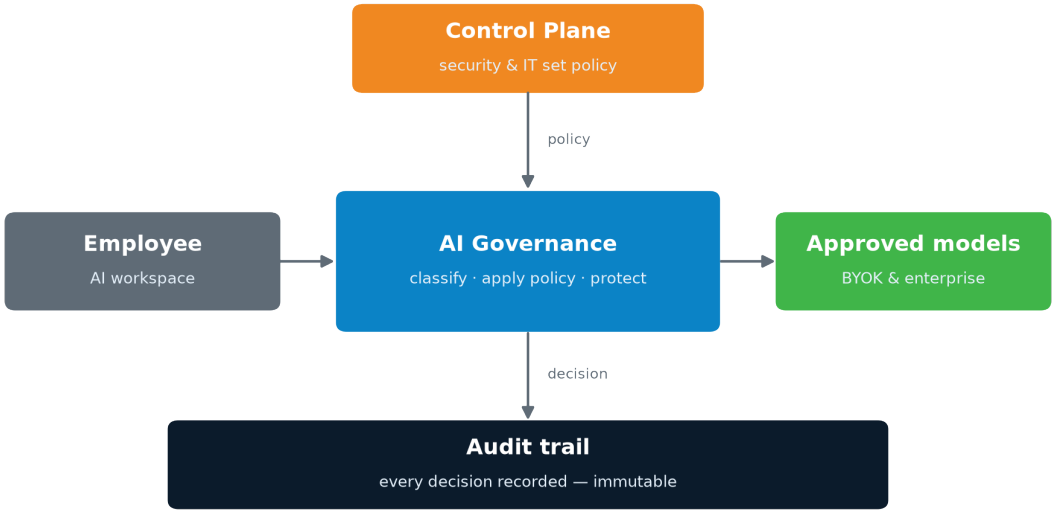
3.6 Show the decision first — The user sees what happened — allowed, redacted, routed, or blocked — and why, before the answer appears. Run in monitor mode to measure exposure first, then switch to enforce with confidence.

4. The system of record: immutable, tamper-evident audit

Every request produces one immutable decision record — who asked, the data class, the action taken, the destination used, and what was redacted — chained with hashes so tampering is detectable. ThreatLens stores hashes and redacted excerpts, never raw sensitive payloads. Governance events stream to your SIEM. This is the evidence you hand to audit and regulators.

5. Architecture and deployment

At a logical level the model is simple: every request flows through the governance layer before it reaches an approved model, the Control Plane sets the policy that governs it, and every decision is recorded.



Every request flows through the governance layer before it reaches an approved model. The Control Plane sets policy; every decision is recorded to the audit trail.

ThreatLens deploys as managed SaaS, in your private cloud, or fully self-hosted — bring your own models and keys (Azure OpenAI, AWS Bedrock). The backend stays private behind a single, rate-limited front door.

6. Governance and access control

RBAC — 21 fine-grained permissions across users, providers, policy, exceptions, incidents, and audit, bundled into roles (Tenant Owner, Security Admin, Policy Manager, Approver, Auditor, User). Least-privilege by default.

SSO / SAML — Microsoft Entra, Okta, and Google, with group→role mapping. Enterprise identity is the source of truth.

Secrets vault — provider and SIEM secrets, local or Azure Key Vault backed.

Self-auditing — every control-plane change is itself recorded in the audit trail.

7. The adoption path

1. **Connect** your models (BYOK) and identity (SSO).
2. **Observe** — turn on monitor mode and measure real exposure across data classes.
3. **Tune** the policy matrix to your risk appetite.
4. **Enforce** — switch enforcement on with confidence.
5. **Roll out** the workspace to employees and review the audit trail.

8. Conclusion

AI adoption is not optional; ungoverned adoption is. ThreatLens makes the safe path the easy path — a single intelligence layer that lets enterprises say **yes** to AI while keeping data, policy, and the record firmly under their own control.

ABOUT THREATLENS

ThreatLens — The Intelligence Layer for Enterprises.

Learn more at thethreatlens.com · Documentation at docs.thethreatlens.com