Aegix

Enterprise AI Privacy (Whitepaper)

Privacy-First Governance Layer for Enterprise LLM Systems

Publisher: Aegean AI Core
Audience: Enterprise security, AI governance, solution engineering
Version: 1.0
Format: Website whitepaper / downloadable PDF

Executive Summary

Large language models are rapidly entering enterprise workflows, but direct interaction with these systems can expose sensitive personal, financial, contractual, and proprietary business data. Traditional privacy tools often focus on isolated pattern matching and do not provide a policy-driven, explainable, enterprise-aware decision layer for AI usage.

Aegix addresses this gap by acting as a **privacy-first AI gateway** between enterprise users, enterprise data, and large language models. It detects sensitive content, sanitizes prompts before model interaction, enforces restoration policies, and records explainable privacy decisions through a **Privacy Decision Graph (PDG)**. The result is a controlled, auditable, and enterprise-aware architecture for adopting AI systems without sacrificing governance.

> Aegix is not just a detector. It is a structured privacy-control and governance layer that enables safe enterprise interaction with LLMs through policy enforcement, explainability, and controlled adaptation.

Why Enterprise AI Needs a Privacy Governance Layer

Enterprises face a practical dilemma: they want to use AI systems to improve productivity, but they cannot allow confidential data to flow into model interactions without strong controls. This challenge is especially visible in use cases involving HR records, contracts, customer data, finance, R&D documents, and internal operational systems.

Three issues repeatedly appear in enterprise deployments:

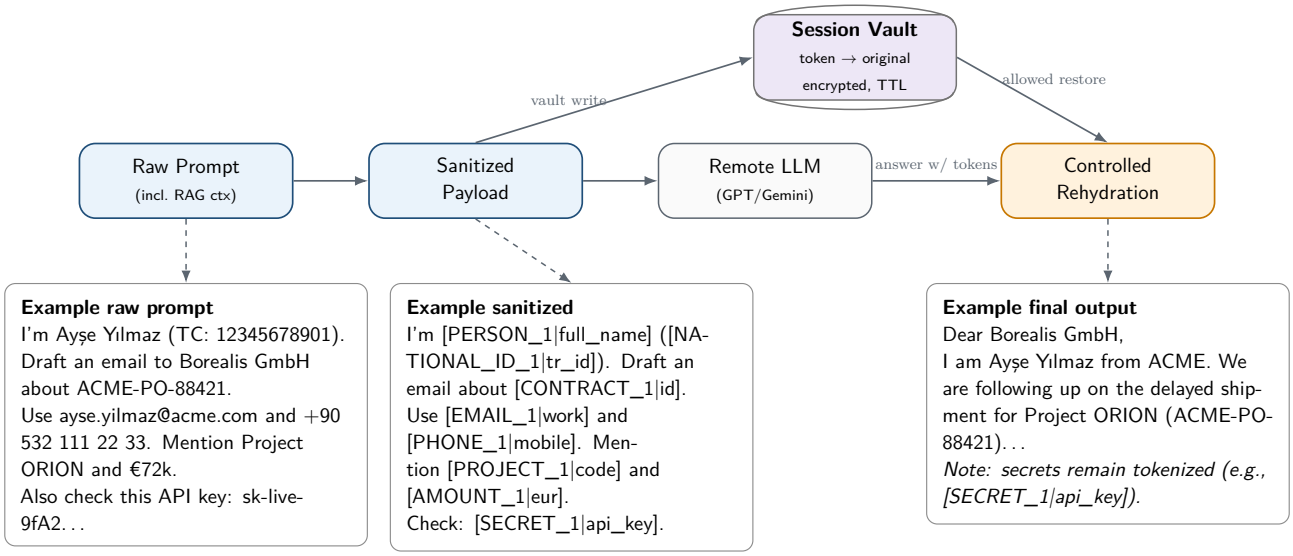
1. **Sensitive data exposure:** prompts may contain personal information, business identifiers, confidential project names, or secrets.
2. **Lack of controllable restoration:** once data is masked, organizations still need a governed mechanism to decide when selected values may be restored.
3. **Limited explainability:** many systems do not offer a clear record of why a value was sanitized, blocked, or restored.

Aegix is designed to solve exactly these problems.

How Aegix Works

Aegix sits between enterprise users and AI models. Its core operating principle is simple: **only sanitized content is sent to the model**. Sensitive values are replaced with structured tokens,

stored securely in a session vault, and restored only when policy and authorization rules permit it.



This architecture shows the essential Aegix promise: enterprise-sensitive content is abstracted before model interaction, governed during processing, and restored only under controlled conditions.

Key Technologies and Product Strengths

1. Policy-Driven Privacy Enforcement

Aegix uses explicit privacy policies to define how prompts and outputs should be handled. Policies define blocked classes, allowed restore classes, invariants, and optional authorization logic by role, department, or group.

This enables organizations to move from informal privacy expectations to **formal, enforceable privacy behavior**. Policy validation can also detect structural contradictions before deployment.

2. Enterprise-Aware Type Registry

Aegix organizes protected information into structured types such as PII, corporate identifiers, financial classes, and secrets. Enterprises can extend this registry with their own business-specific types, allowing protection logic to reflect real organizational semantics rather than generic consumer patterns.

3. Dictionary and Schema-Enrichment Support

Enterprise data rarely appears only in generic forms. Aegix supports:

- dictionary-based detection for known high-value identifiers
- schema enrichment to map enterprise fields and structures to privacy types
- optional LLM-assisted type suggestion generation, with local-first deployment options

This gives Aegix a strong advantage in domains where internal project names, financial fields, HR entities, or contract identifiers matter.

4. Privacy Decision Graph (PDG)

Aegix records explainable privacy decisions in a **Privacy Decision Graph**. The PDG makes it possible to inspect:

- what was detected

- which policy rules were evaluated
- why rehydration was allowed or denied
- how risk evolved across a session

This supports debugging, auditability, compliance discussions, and internal governance.

5. Controlled Calibration and Overlay Rules

Aegix supports structured administrative feedback through scenario-based calibration. Feedback produces candidate rules, which are validated and published as overlay rules on top of the base policy.

This allows organizations to refine behavior over time without losing transparency or weakening core privacy guarantees.

6. Learning Pipeline and AI-Assisted Risk Scoring

Aegix can collect learning events and PDG snapshots to support advisor models and AI-assisted risk scoring. This enables a future-facing learning pipeline where the system improves contextual risk evaluation while still keeping policy enforcement as the final authority.

Deployment Value for Enterprise Teams

Aegix is designed for real enterprise deployment patterns, including:

- on-premise or controlled cloud deployments
- local-first privacy enforcement
- multi-tenant governance
- integration with remote or local LLMs
- security-conscious administrative workflows

This makes it relevant not only as a developer tool, but as an enterprise AI control layer that can be adopted by security teams, compliance teams, and AI platform teams together.

> Aegix is especially strong in environments where organizations need both **AI enablement** and **governable privacy control**. It reduces the trade-off between using LLMs productively and protecting enterprise-sensitive data.

Conclusion

Enterprise adoption of LLMs requires more than model access. It requires a privacy and governance layer that can detect sensitive data, enforce rules, explain decisions, and adapt safely over time.

Aegix delivers this through a unified architecture built around policy enforcement, structured typing, enterprise-aware detection, explainable decision graphs, and controlled learning. The result is a practical and defensible way for organizations to adopt AI systems while keeping privacy, security, and governance under control.