aligne.ai

WHITE PAPER · 2026 · RUNTIME AI GOVERNANCE

Your AI Policies Won't Save You.

Why governance that lives in documents is failing organisations that run AI in production and what the runtime control layer that actually works looks like.

aligne.ai · In partnership with IBM · 2026

Gurpreet Dhindsa
Director, Product & Responsible AI, Aligne AI

Amit Joshi
Managing Director, Aligne AI

EXECUTIVE SUMMARY

Your AI Policies Won't Save You.

Many organisations believe their AI risks are adequately managed through policies, governance frameworks, vendor assessments, and legal reviews. While these controls remain essential, they provide only a static view of risk in an environment that is increasingly dynamic.

The critical challenge is not whether governance policies exist, it is whether AI systems are operating within those policies in practice. Traditional governance approaches offer limited visibility into how generative AI is being used across the organisation, what data is being shared, which models are being accessed, how autonomous workflows are making decisions, and whether established controls are being consistently enforced.

> **No policy document has ever blocked a hallucination, an unauthorised disclosure, or a model behaving outside its mandate**

As AI adoption accelerates, governance must evolve from documentation and oversight to continuous operational control. Runtime governance provides real-time visibility, monitoring, and enforcement across AI interactions, enabling organisations to detect policy violations, manage emerging risks, and demonstrate compliance with confidence.

This paper argues that policy-based governance alone is no longer sufficient for organisations deploying AI at scale. It explores the role of runtime governance, the risks associated with unmanaged AI operations, and the practical steps required to establish auditable, enterprise-grade AI governance capabilities.

The numbers confirm what the argument implies: adoption has outpaced oversight, and the gap is widening.

$37B
enterprise AI spend in 2025
Menlo Ventures, 2025

31%
of AI use cases now in full production
ISG, 2025

78%
of executives cannot pass an AI governance audit in 90 days
Grant Thornton, 2026

THE CORE PROBLEM

Governance That Lives in Documents Cannot Enforce Itself

Every serious organisation has AI governance. Nearly all of it is the same thing: a set of rules written down somewhere that describes what AI should and should not do.

These documents are necessary. They are not sufficient. They answer one question: should this AI system be allowed to exist? and stop there. They do not answer what happens after it goes live.

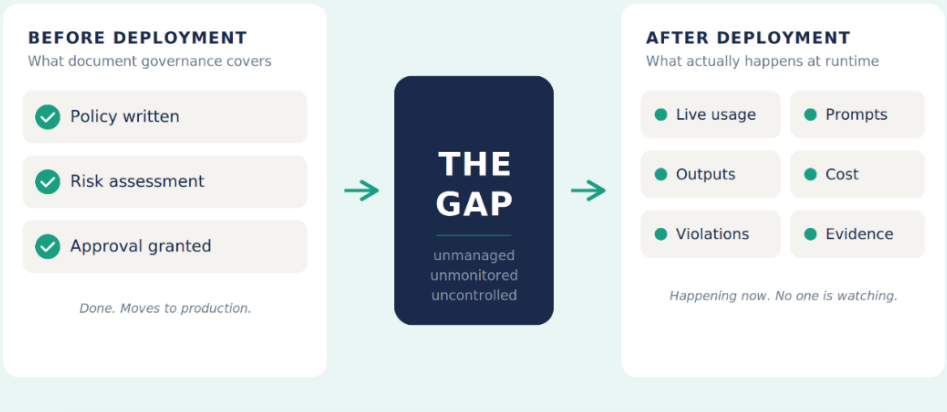
What policy governance cannot tell you

- Whether sensitive data appeared in a prompt this morning
- Which model your internal tools actually called and whether it was approved
- What an AI agent did autonomously between 2am and 6am while no-one was watching
- Whether the output your customer-facing assistant gave last week was accurate, appropriate and compliant
- What your AI deployment is costing by team, by use case and by model
- What evidence exists of your governance controls working before an auditor asks for it

> **No policy has ever intercepted a prompt containing client data, redacted it before it reached the model, and logged the violation — all within the same second. That is not what policies do. That is what runtime governance does**

Policy governance ends at the point of deployment. Everything that happens after that falls into an operational gap that no written policy can close. Figure 1 illustrates where that gap sits.

Figure 1: The governance gap



Before deployment: policy, assessment, approval. After deployment: live usage, prompts, outputs, cost, violations and evidence.
The gap between the two is where the risk lives.

The practical consequence of this gap becomes clear when you compare what each governance model actually delivers. The table below sets out the distinction between pre-deployment governance and runtime governance — not as competing approaches, but as two layers that serve fundamentally different purposes.

Pre-deployment governance	Runtime governance
Defines what should happen	Monitors what is happening
Reviews use cases before deployment	Observes systems during use
Sets policies and standards	Enforces policies in live interactions
Documents intended controls	Captures evidence of actual controls
Reviews risk periodically	Tracks risk continuously
Focuses on approval	Focuses on operation

RUNTIME GOVERNANCE

What Is Runtime Governance?

Runtime governance is the ability to monitor, control, optimise and evidence AI system behaviour while the system is being used.

It operates as a persistent control layer sitting between your users, your applications, your data and your models. Every prompt submitted to a model, every response returned, and every action taken by an autonomous agent passes through this layer where it is intercepted, checked against defined policy, logged, and preserved as auditable evidence. The interaction between humans and AI systems does not occur outside governance. It occurs within it.

It is not a replacement for policy-level governance. The two operate at different points in the AI lifecycle. Policy-level governance determines whether a system should be deployed. Runtime governance determines what happens during deployment, continuously and at scale. Organisations that rely on the former without the latter have defined their intent but have no means of verifying whether that intent is being observed.

Runtime governance addresses this through four interconnected capabilities, each operating in real time across every AI interaction the organisation makes.

> **Pre-deployment governance answers: Should this AI system go live? Runtime governance answers: What happens after it does?**

RUNTIME RISKS

The Risks You Missed Don't Appear in Policy Reviews. They appear at runtime.

Policy governance operates at the point of approval. What happens after that approval is granted falls entirely outside its view.

The data leakage does not occur in the risk assessment. It occurs the moment a user pastes client data into a prompt. The prompt injection does not occur in the security review. It occurs during a live interaction with a model that has no awareness of your policies. The cost overrun does not appear in the approved budget. It accumulates quietly as token usage expands across teams, workflows and agents that nobody accounted for when the project was signed off.

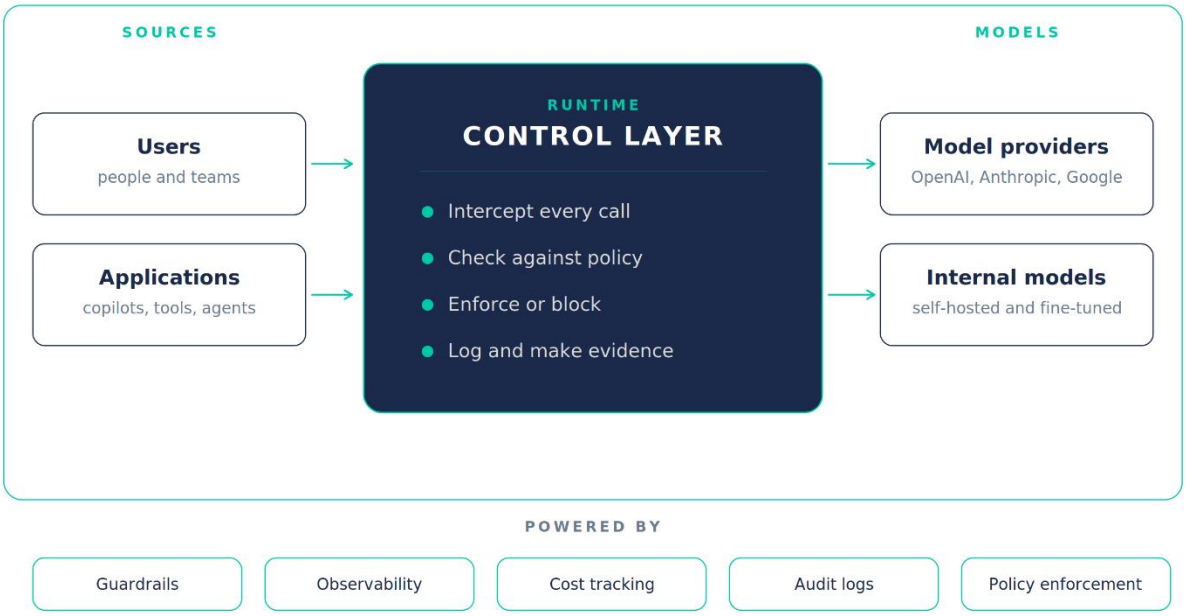
Policy governance captures none of this. It cannot, because it is not present when AI is actually being used. The gap between what governance documents describe and what AI systems actually do is not a flaw in the writing. It is a structural limitation of the approach. The table below sets out exactly where that limitation falls and what runtime governance does instead.

Runtime risk	How it manifests	Runtime governance control
Prompt injection	Malicious inputs manipulate model behaviour and may bypass safety measures (OWASP LLM01:2025)	Input scanning, prompt analysis, policy checks, real-time blocking
Sensitive data leakage	Users paste PII, PHI or commercially sensitive data into external AI tools	PII/PHI detection, redaction at inference, blocking, immutable audit log
Unsafe outputs	Model returns harmful, inaccurate or policy-violating content	Response checks, content guardrails, escalation, human-in-the-loop triggers
Unapproved model use	Teams route requests to models not approved by security or compliance	Gateway-based access control, model allowlisting, routing enforcement
Cost overruns	Token usage expands uncontrolled across teams, agents and workflows	Usage tracking, budget alerts, model optimisation routing, attribution by team
Agentic overreach	Autonomous agents act beyond their defined scope without oversight	Permissioning controls, tool restrictions, human approval gates, continuous monitoring

Runtime risk	How it manifests	Runtime governance control
Audit gaps	No evidence exists of what AI systems did and whether controls were applied	Interaction logs, metadata capture, automated compliance reporting

Runtime governance is where AI risk management becomes operational.

Figure 2: The runtime governance layer



Every prompt and response routes through one control layer before reaching a model.

AGENTIC AI

Agents Change Everything.

Copilots answer questions. Agents take actions.

The distinction between copilots and agents is not a matter of degree. It is a fundamental shift in how AI operates within an organisation.

Copilot-style tools function as sophisticated assistants. A user submits a prompt, the model generates a response, and a human evaluates the output before deciding whether and how to act on it. The human remains in the loop at every stage, and the blast radius of a poor response is limited to the quality of a single answer.

Agentic AI operates on an entirely different basis. An agent does not wait for instructions at each step. It decomposes a goal into a sequence of actions, retrieves information from live data sources, calls external tools and APIs, interacts with internal systems, triggers downstream workflows, updates records and sends communications all autonomously, and often through permissions that were granted without adequate consideration of what an autonomous system might do with them. In many organisations, this is already happening without meaningful oversight.

The scale of agentic deployment has already outpaced the governance response.

37
average AI agents deployed per enterprise
Security Boulevard, 2026

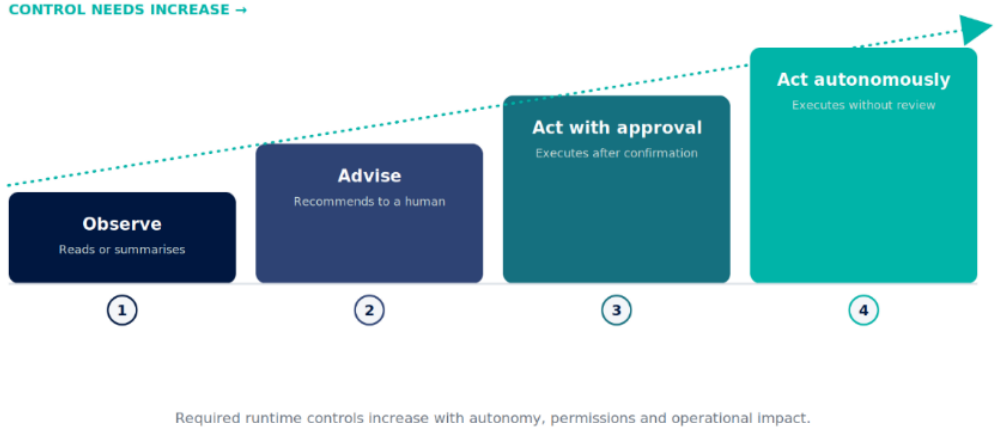
>50%
of those agents run with no security oversight or governance
Security Boulevard, 2026

1,445%
increase in multi-agent system deployments, 2024–2025
Industry data, 2026

Not all agents carry the same risk. The level of governance required scales directly with the degree of autonomy the agent has been granted — from agents that read and summarise information through to those that execute actions, trigger workflows and operate without human review. Figure 3 illustrates this spectrum.

The practical implication is that each level of autonomy demands a different set of runtime controls. The table below sets out the minimum governance requirements at each stage. Organisations should identify where their current agents sit on this spectrum and assess whether the controls in place are proportionate to the autonomy they have been granted.

Figure 3: Agent autonomy and control needs



Required runtime controls increase with autonomy, permissions and operational impact.

Agent level	What it does	Minimum runtime controls required
Observe	Reads or summarises information	Logging, data access control, output review
Advise	Recommends actions to a human	Traceability, rationale capture, human decision point
Act with approval	Executes actions after human confirmation	Approval workflow, permission checks, action logs
Act autonomously	Executes within defined boundaries without human review	Continuous monitoring, guardrails, rollback capability, incident response

The bottom two rows of this table represent where most organisations are currently weakest and where the consequences of unmanaged autonomy are most severe.

> **A pre-deployment assessment tells you what an agent is designed to do. Runtime governance tells you what it actually did. For autonomous systems, only one of these matters.**

WHY POLICY-BASED GOVERNANCE FALLS SHORT

Why Policy-Based Governance Is No Longer Enough

Policy-based governance was designed for a world where AI was reviewed before deployment and rarely changed after. That world no longer exists. Organisations still relying on it are not just behind, they are exposed in ways their governance frameworks cannot detect.

1. The enforcement gap is structural not a training problem

Every AI policy has the same weakness: it was written at a point in time, by people who could not anticipate every model update, every new feature, every edge case that would emerge in production. The moment a tool is updated, a new model is deployed, or a workflow is extended, the policy is already behind. Policy-based governance does not fail because people ignore it. It fails because it cannot keep pace with the rate at which AI changes.

And even when the policy is current, enforcement depends entirely on humans remembering to follow it across hundreds of daily AI interactions, under deadline pressure, in tools that make it frictionless to do the wrong thing. That is not a behavioural problem. It is a structural one. Runtime governance removes the dependency on human memory by enforcing controls automatically at the point of inference, every time, regardless of who is using the system or under what pressure.

2. When a regulator asks, a policy is not an answer

The FCA's SM&CR requires senior managers to demonstrate accountability for AI oversight, not merely assert it. The EU AI Act mandates automated logging and audit trails for high-risk systems. When a regulator asks what your AI did, what it returned, and whether controls were applied, a policy document is not a response. The absence of runtime logs is the absence of governance.

Runtime governance produces the evidence regulators need: immutable interaction logs, violation records, model call metadata, and automated compliance reports, ready before anyone asks. Organisations with this capability walk into scrutiny with confidence. Those without are reconstructing events from incomplete records, under pressure, after the fact.

3. Agents are not copilots and governance that treats them the same has already failed

Policy-based governance assumed a simple model: a human asks, an AI responds, a human decides. Agentic AI breaks every assumption in that model. Agents decompose goals into sequences of actions, retrieve live data, call external tools, update records, trigger workflows, and operate through permissions granted without adequate review, often without anyone watching.

A policy review cannot anticipate what a chain of autonomous actions will produce at 3am. Only continuous runtime monitoring can. Organisations deploying agents without runtime controls are not just ungoverned, they carry real-world exposure their risk frameworks have not accounted for.

4. False confidence is the most dangerous outcome of all

The most dangerous thing about policy-based governance is not what it misses, it is the confidence it creates. Organisations that have completed a risk assessment, reviewed vendor contracts, and published an AI policy often genuinely believe they are in control. That belief is not tested until something fails: a regulatory inquiry, a data incident, an agent acting out of scope. By the time the gap is visible, the question is no longer how to prevent the problem. It is how to explain it.

The same blind spot applies to cost. Without runtime visibility, AI spend is unattributable: no one can tell you what each team is spending, which use cases are driving token consumption, or where budget is being wasted. Organisations routinely discover their AI costs are two or three times what they expected, with no way to trace where the money went.

Runtime governance replaces assumed assurance with provable assurance. It makes risk visible, cost attributable, and evidence available, not because the organisation trusts its people less, but because trust alone is no longer sufficient at the speed and scale at which AI now operates. The organisations that understand this are already building the control layer. The ones that do not will discover the gap when they can least afford to.

COST GOVERNANCE

Runtime Governance Is Also Cost Governance.

Most organisations frame runtime governance as a risk and compliance question. It is also a financial discipline question and for many organisations, the cost case is equally compelling.

GenAI cost does not behave like traditional software cost. It scales with usage, not licences. It is driven by prompt volume, response length, token consumption, model selection, context window size, tool calls, agent loops and repeated queries. Without visibility into these drivers, organisations cannot manage them.

The scale of the problem is already measurable. The proportion of organisations actively managing AI spend doubled in the past year alone, yet most still lack the infrastructure to attribute cost at a meaningful level of granularity. The data below illustrates both the pace of change and the size of the opportunity that runtime cost controls can unlock.

63%

of organisations actively managing AI spend in 2025 — up from 31% the previous year

FinOps Foundation, 2025

18–20%

higher cost of external model calls versus optimised internal routing

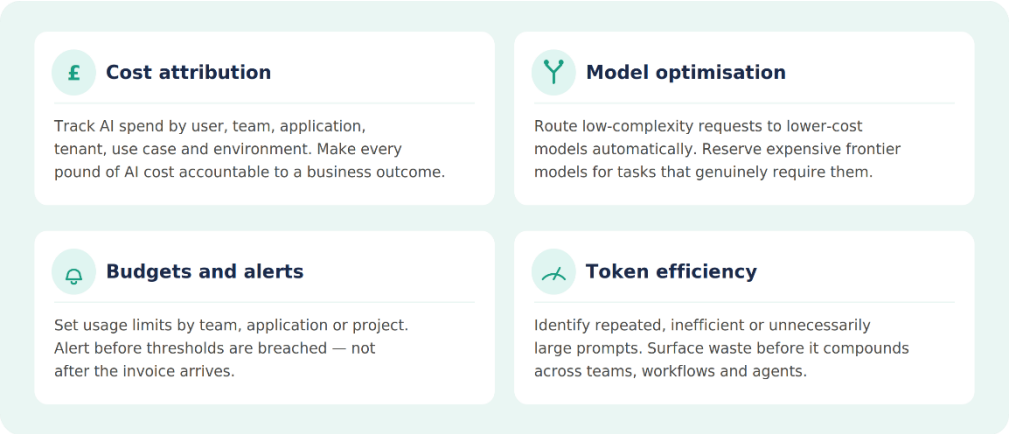
Industry benchmark

40–60%

potential token cost reduction through smart model routing and caching

AWS Bedrock data

Runtime cost governance operates across four disciplines, each addressing a different dimension of the problem. Together they provide the visibility, attribution and control required to treat AI expenditure as a managed line item rather than an uncontrolled variable.



Uncontrolled GenAI usage does not only create risk exposure. It creates margin exposure. Runtime governance turns AI cost from an unmanaged variable into a controlled business metric.

BOARD-LEVEL QUESTION

The Question Boards Are About to Start Asking

The FCA has confirmed that senior managers will be personally accountable for AI harm under SM&CR. The EU AI Act mandates automated logging and audit trails for high-risk AI systems from August 2026. The next phase of regulatory enforcement will not ask whether you have AI policies. It will ask whether you can prove your AI is under control.

The organisations that will answer that question with confidence are not the ones with the most policies. They are the ones that have already built runtime governance into their AI infrastructure — not as a compliance overlay, but as a core operational capability. They know where their AI is running, what it is doing, and what it is costing. They can demonstrate control before anyone asks.

The organisations that will struggle are those that assumed policies were enough — and discovered they were not when it was too late to fix it quickly.

> **The real test of AI governance is not whether your organisation has written rules. The real test is whether it can see, control, optimise and evidence AI usage at the moment it happens.**

That is not a future aspiration. It is a present requirement. Connect with Aligne to build the capability that closes the gap

Join the conversation.

Aligne AI works with leadership and technical teams to identify runtime governance gaps, prioritise where risk and cost exposure are highest, and build a clear path from current state to audit-ready operations.

To start the conversation, contact Sadiq Merali at sadiq.merali@aligne.ai

About Aligne AI

Aligne AI specialises in practical AI governance for regulated organisations. Built on 100+ years of combined GRC leadership experience. IBM Gold Partner in EMEA. aligne.ai · info@aligne.ai

About Altrum

Altrum is Aligne AI's runtime governance platform the control layer for enterprise AI. Monitor, control, optimise and evidence AI system behaviour at the point of inference. No-code policy builder. Real-time enforcement. Audit-ready reporting.

References

Menlo Ventures (2025). The State of Generative AI in the Enterprise. menlovc.com

OpenAI (2025). The State of Enterprise AI 2025.

ISG (2025). State of Enterprise AI Adoption Report. isg-one.com

OWASP GenAI Security Project (2025). LLM01:2025 Prompt Injection. genai.owasp.org

NIST (2024). AI Risk Management Framework and Generative AI Profile. nist.gov

Microsoft (2025). Agentic AI Security and Governance. learn.microsoft.com

FinOps Foundation (2025). State of FinOps Report. data.finops.org

Security Boulevard (2026). Enterprise AI Agent Deployment Statistics.

Grant Thornton (2026). Enterprise AI Governance Readiness Survey.

HM Treasury Select Committee (2026). AI in Financial Services Report.

This document is intended as a thought leadership and campaign asset. It should not be treated as legal, regulatory or security advice.