

★ BOARD-LEVEL GOVERNANCE

Governing the Autonomous Organization

Why AI governance is the defining enterprise discipline of 2026

AUTHOR

Peter Sondergaard

Chairman of the Board, 2021.AI
Strategic Advisor, Polaris I/O
Former Executive Vice President, Gartner

IN BRIEF

Every governance failure of 2026 shares one root cause: leadership could not see what its own AI was doing until the bill, the breach, or the regulator arrived. This paper sets out the cost, a five-stage maturity model, and the operating discipline that closes the gap.

Contents

03	Executive summary
04	The cost of getting this wrong
05	The scale problem: why this is accelerating, not stabilizing
07	What's actually working
07	How governance actually starts
09	A practical maturity model
10	See it. Spot it. Steer it. Governance before the incident report
11	Where Polaris I/O fits

Enterprise AI spending in the US reached roughly \$11.6 million per company in 2026, yet 56 percent of CEOs report no measurable revenue or cost benefits from it. Token consumption, the literal currency of AI, is projected by Goldman Sachs to grow 24x by 2030 as agentic systems shift from answering questions to running unattended, around the clock.

Data and AI remain the "oil and combustion engine" of the transformation toward an autonomous-driven business model and economy. Against this backdrop, organizations are trying to navigate classic technology adoption issues, such as AI budgets, investment ROI, leadership capability requirements, and organizational changes, while strategically driving AI adoption at scale. They are navigating the real and growing chasm between AI innovation and AI adoption at scale.

The common thread among winners and losers is not model quality. It is a combination of classic challenges organizations face when deploying transformative technologies. These include talent, leadership capabilities, organizational transformation, and governance. Of these, governance is critical. It is about who is accountable, what is measured, and whether the organization can see what its own AI is doing before a competitor, investor, employee group, customer, regulator, journalist, or plaintiff's attorney.

This paper explains why AI governance has shifted from a compliance checkbox to a board-level fiduciary issue, the cost of getting it wrong, what a practical maturity model looks like, and how organizations are building the muscle to govern AI at enterprise scale.



\$11.6M

AVG. AI SPEND
PER COMPANY, 2026



56%

CEOs REPORT NO
MEASURABLE
REVENUE OR COST BENEFIT



24x

PROJECTED TOKEN GROWTH
BY 2030

01 The cost of getting this wrong

Governance failures in 2026 are not theoretical, and they are increasingly not about a model producing a wrong answer. They are about enterprises losing visibility into their own AI usage until the bill, the breach, or the regulator arrives.

IN THE NEWS • TWO 2026 GOVERNANCE FAILURES IN REAL TIME

Meta built a culture problem it could not see coming, then had to publicly reverse it.

In late 2025, Meta made “AI-driven impact” a core performance expectation for employees, with bonuses tied to it¹. An employee built an internal leaderboard called “Claudeconomics,” turning that incentive into a competition, ranking all 85,000-plus employees by token consumption, with titles like “Token Legend” and “Session Immortal.”² Within a single 30-day window, employees burned through 73.7 trillion tokens, some running AI agents idly just to climb the rankings.³ Meta’s own CTO, Andrew Bosworth, was forced to walk it back in an April 2026 memo: “Nobody should be using AI just for the sake of using AI ... token usage itself is not a measure of impact.”⁴ Meta shut the leaderboard down and is now building a centralized monitoring system, called “AI Gateway,” with real-time budgets and automated alerts for unusual spending⁵ — the exact governance layer that should have existed before the incentive program launched, not after.

Uber and Microsoft discovered their 2026 budgets were obsolete within months.

Uber’s CTO, Praveen Neppalli Naga, said publicly in 2026: “I’m back to the drawing board, because the budget I thought I would need is blown away already.”⁶ Engineering adoption of agentic coding tools jumped from 32 percent to 84 percent of Uber’s roughly 5,000-engineer organization between December 2025 and March 2026; by April, the entire annual AI coding budget was gone, with monthly per-engineer API costs running \$500 to \$2,000.⁷ Uber’s COO later acknowledged the harder problem: “It’s very hard to draw a line between one of those stats and, ‘Okay, now we’re actually producing 25 percent more useful consumer features.’”⁸ Microsoft canceled most of its internal Claude Code licenses months after rolling them out, citing unsustainable token bills.⁹ ServiceNow and Priceline ran into the same wall, exhausting tool budgets or blowing past cloud credit allocations within the first months of the year.¹⁰

The root cause is structural, not a vendor pricing problem. Per-token prices have actually fallen roughly 67 percent year over year.¹¹ The bills are rising anyway, because agentic workflows consume 5 to 30 times more tokens per task than a standard chatbot query,¹² and most 2026 budgets were built in late 2025, before agentic

tools existed at scale. The result of lacking AI budget control will be that projects get canceled due to cost overruns alone — not technical failure, not poor fit; pure economics that nobody governed in advance.

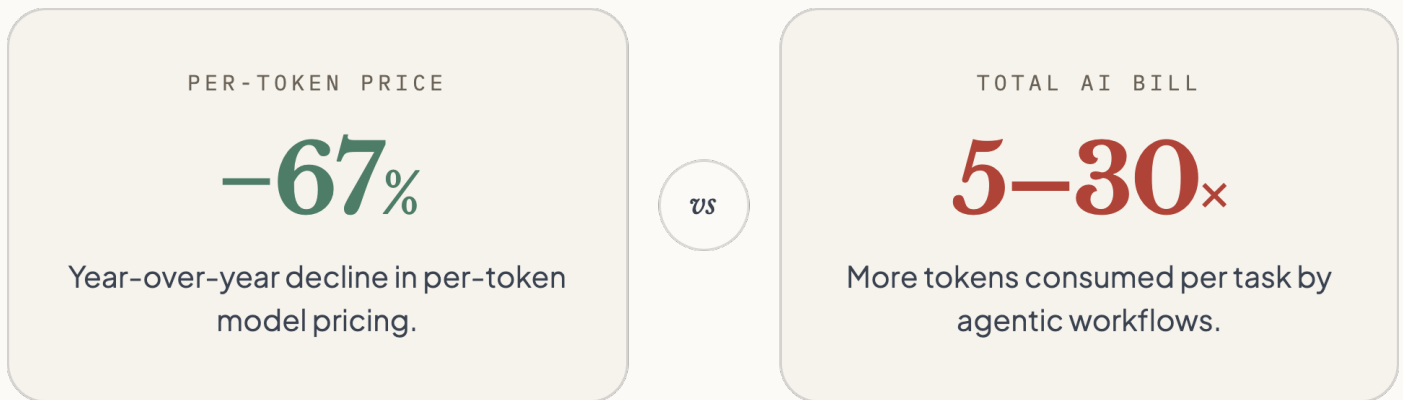


Figure 1 · Most 2026 budgets were built in late 2025, before agentic tools existed at scale — which is why the bill is rising even as unit prices fall.

None of these governance failures originated in a model producing a bad output. Every one of them originated in the same gap: leadership had no real-time visibility into how AI was actually being used, what it was costing, or what it was exposed to, until it was already too late to manage cheaply. AI cost becomes a governance issue, necessitating greater, distributed governance aligned with the financial management structure of each organization.

02 The scale problem: why this is accelerating, not stabilizing

Organizations are scaling AI usage, but not just through large, monolithic platforms that embed AI; it is happening more through a mushrooming number of AI implementations. Scale is not monolithic — it is distributed. Scaling AI becomes increasingly challenging, requiring a sophisticated governance platform to manage AI. Three converging trends are why governance cannot wait for a calmer moment.

Token growth is compounding.

Goldman Sachs' 2026 research projects global token consumption will grow 24x by 2030, reaching roughly 120 quadrillion tokens per month, driven not by more people asking more questions, but by agents that never stop working.¹³ By 2040, enterprise agents are projected to account for over 70 percent of all token usage.¹⁴

GLOBAL TOKEN CONSUMPTION PER MONTH

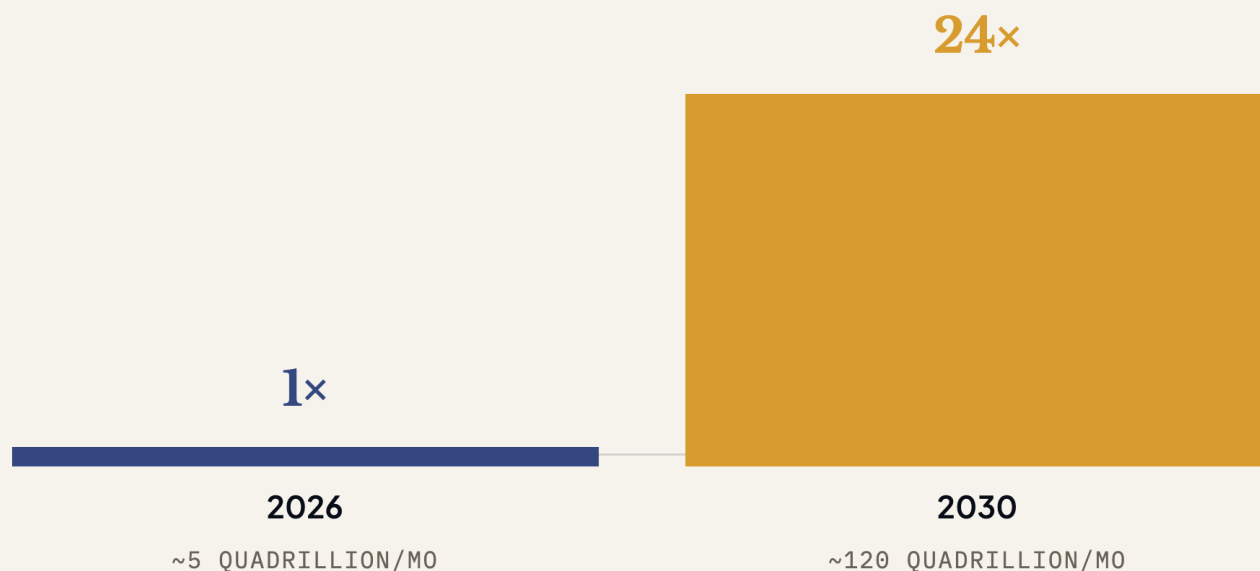


Figure 2 · By 2040, enterprise agents are projected to account for over 70 percent of all token usage. Source: Goldman Sachs, 2026.

Shadow AI is the default state, not the exception.

I have tracked this pattern since 2019: “It’s highly likely that we’ll witness a surge in the deployment of ‘Shadow AI’ models, not unlike the historical spread of Shadow IT ... This widespread adoption will cause the proliferation of a shadow AI model environment.” My research at the time estimated the average large enterprise would go from roughly five deployed AI models to 35 within four years — a figure that, if anything, understated what agentic AI has since done to model proliferation.

Organizational and business process reinvention.

As I said in a 2025 interview: companies are running into a productivity ceiling not because the technology underperforms, but because “AI doesn’t deliver any productivity benefit until we rework the business process and how tasks are composed around jobs.”¹⁵ The capability exists; the organization has not yet restructured its work to use it.

The combined effect of these three trends forces organizations to rethink their approach to managing — and governing — AI.

03 What's actually working

It is not all bleak. The data also shows a clear, learnable pattern among organizations that are succeeding in their AI efforts.



Buy and partner, do not build alone.

Five years ago, most large organizations built their own AI environments, including AI governance platforms. Today, successful organizations purchase AI tools from specialized vendors and partner with implementation providers with demonstrable expertise in scaling AI.



Managing data, structured or unstructured, is key to scaling.

Most organizations have data scattered throughout systems, platforms, and the organization. Strategically managing data across the organization has proven impactful for successful, scalable AI projects.



Narrow, back-office use cases outperform flashy front-office ones.

Document processing, procurement, risk management, and contract management consistently deliver clearer, faster ROI than broad customer-facing deployments.



Distributed accountability beats centralized experimentation.

Successful organizations push experimentation to the team level while holding a board-anchored governance charter above it. The governance owner is singular; the experimentation is not.



One named, accountable owner beats a committee.

Across multiple 2026 governance failure analyses, the single most consistent fix recommended by risk professionals is naming one person, not a working group, who coordinates (not necessarily owns) AI governance and can be asked directly what went wrong.

04 How governance actually starts

My own framework is a useful and well-tested starting point. Governance has to answer three questions most organizations skip.

1

What does AI governance mean here, specifically?

Not the public policy definition — the operational one: what do explainable, transparent, and ethical mean for your company's data, models, and decisions?

2 Who is accountable, by name and function?

My model assigns ultimate accountability to the CEO, audit responsibility to the board (specifically the audit committee), legal and risk exposure to the General Counsel, financial risk to the CFO, and ongoing operational coordination to a Chief Data Officer or equivalent role. While implementations may vary by organization, it is paramount that organizations explicitly address this now.



Figure 3 · Accountability map. One name per function — not a shared responsibility model.

3 How is it measured?

What you can't measure, you can't manage. My recommended KPI categories: data lineage and quality, AI security, cost and value, bias, accountability (who used what system, for what decision, when), audit trail continuity, and time-based impact tracking.¹⁶

Crucially, this needs to be built into the AI environment itself as a live operating layer — not bolted on as a quarterly compliance review.

05 A practical maturity model

Most maturity models describe where a company sits. Few tell leadership what to actually do about it. Drawing on what I have seen across enterprises wrestling with this problem — and on my own operating concept of compliance debt, the accumulation of outdated policies, incomplete risk assessments, and unaddressed governance gaps that silently build until they force a costly, all-at-once reckoning — the five-stage model below pairs each stage with the small set of moves that actually advance a company to the next one.

STAGE	GOVERNANCE REALITY	TOP PRIORITIES TO REACH THE NEXT STAGE	WHERE COMPANIES SIT
1 Invisible	No registry of AI in use. Governance is a policy document nobody operationalizes. Compliance debt accumulates invisibly.	Build a single source of truth for what AI is actually running. Name one accountable owner. Put a dollar figure on current token and tool spend.	Most private or public organizations
2 Reactive	Shadow AI, including unmanaged token costs, is the dominant mode of adoption. IT and Legal discover implementations after the fact, not before.	Bring shadow usage into the light without punishing the teams who found value first. Stand up a lightweight intake process. Set a budget ceiling with real alerts, not a year-end review.	The majority of mid-market and many large enterprises today
3 Structured	A named governance owner exists. A cross-functional committee (Legal, IT, Compliance, Data) is active. KPIs are defined but not yet automated. An AI governance software layer is being implemented.	Automate the KPIs still tracked by hand. Connect cost monitoring to the same system that tracks compliance, so one dashboard answers both questions. Define the playbook for what happens when a signal fires, not just who receives it. Plan for governance as part of the AI software stack.	Roughly one in five large organizations
4 Operational	Organizational structures for AI governance are in place; the organization has an AI governance software layer supporting all new AI solutions.	Push governance ownership down to the team level without losing the central view. Extend monitoring outside the company — vendors, regulators, competitors — not just internal usage. Start treating governance fluency as a hiring, promotion, and supplier evaluation criterion, not just a compliance one.	Early-mover enterprises
5 Embedded	Governance is built into hiring criteria, performance standards, and agentic workflow design from day one — not managed by a separate function alongside the business. AI governance is a mature software layer in the organization.	Stay here. The work at this stage is keeping pace with how fast agentic capability is changing, not solving a fixed problem once.	The stated direction of the market, not yet broadly achieved

06 See it. Spot it. Steer it.

Most governance programs fail for a simple reason: they are built to investigate what already happened, not to catch what is about to. By the time Legal or Compliance opens a file, the cost has usually already been incurred, the budget already blown, the regulator already notified. The discipline that actually prevents a severe governance incident has to operate on the same three-part logic regardless of what you are protecting: see everything, spot what matters, steer before it escalates.



STEP ONE

See it

Total visibility into what is actually happening.



STEP TWO

Spot it

Filtering thousands of changes down to the few that matter.



STEP THREE

Steer it

Acting on the signal while it is still cheap to act on.

See it.

Total visibility into what is actually happening, not what the policy says should be happening. Most enterprises cannot answer basic questions about their own AI footprint: how many models are in production, who deployed them, what data they touch, what they cost this month versus last month. That blindness is the precondition for most failures. Meta did not have a budget problem first; it had a visibility problem. The cost crisis was the symptom that showed up only once usage had already exploded.

Spot it.

Translating thousands of small changes into the handful that actually matter. Seeing everything is necessary but not sufficient; most enterprises that do build a model registry or a usage dashboard drown in noise and still miss the signal. What is worth flagging to a governance committee or the CFO is narrower than what is worth logging: a new vendor contract clause that shifts liability, a spike in token consumption that breaks the pattern of the prior quarter, a regulatory filing in a jurisdiction where the company operates, a competitor's publicized AI incident that maps to an identical exposure internally, a model update from a vendor that silently changes behavior in a regulated workflow. Each of these is a pre-intent signal in the same sense Polaris I/O applies the term to revenue and risk: evidence that something material is forming, well before it shows up as an incident, a fine, or a headline. That signal discipline — filtering what is worth a governance committee's attention from everything else that is merely worth logging — is precisely what Polaris I/O is built to provide.

Steer it.

Acting on the signal while it is still cheap to act on. This is the step most governance programs skip entirely, because seeing and spotting feel like the finished product. They are not. Steering means a named owner gets the signal, has the authority to act on it, and has a defined playbook for what acting looks like: pause a workflow, tighten a budget cap, escalate to the audit committee, request a vendor disclosure, retrain a model, or simply document the decision to accept the risk. Uber’s and Meta’s recoveries in 2026 both followed this exact arc once they finally built it — real-time monitoring feeding a specific, accountable response, instead of a quarterly report feeding a committee meeting.

The companies that avoid the next AI catastrophe are not the ones with the most sophisticated models. They are the ones who shortened the distance between seeing a signal and steering a response from months to days.

07 Where Polaris I/O fits

Governance maturity is not built by buying a single tool. It is built by combining three things most enterprises do not have in one place: visibility into what is changing inside the business that creates AI-related risk and opportunity, a proven governance operating framework, and the software to run it continuously. Mapped against see, spot and steer, that means three distinct jobs need to get done — and they rarely live under one roof.

See it VISIBILITY	Instrumentation inside the AI environment itself: a live registry of what is running, who deployed it, what data it touches, and what it costs month over month.
Spot it POLARIS I/O	Pre-intent signal intelligence: the regulatory filings, vendor and contract changes, leadership moves and competitor incidents that become AI exposure — surfaced while the cost of acting is still low.
Steer it OPERATING FRAMEWORK	A governance operating framework with named owners, defined playbooks and the authority to act — running continuously in software, not as a quarterly review.

Figure 4 · Three jobs, rarely under one roof. Polaris I/O operates on the spotting problem specifically.

Polaris I/O, where I also serve as Strategic Advisor, operates one layer earlier — on the spotting problem specifically. Its pre-intent signal intelligence platform was originally built to surface the leadership changes, regulatory shifts, M&A activity, and supplier risk that matter to revenue teams before a competitor acts on them first. The same underlying discipline applies directly to AI governance: catching a regulatory filing in a jurisdiction where the company operates before it becomes a compliance gap, flagging a vendor's model update or contract change before it creates a liability question, or identifying the pattern in a competitor's publicized AI incident before it replicates internally. It is the layer that gets the right signal to the right person while the cost of acting is still low.

Endnotes

- 01 The Information, reporting on internal Meta memo, April 2026.
- 02 The Information; reported widely via MLQ News, Fortune, and The Pragmatic Engineer, April 2026.
- 03 Edgen / The Information, internal Meta dataset reporting, June 2026.
- 04 Andrew Bosworth, internal Meta memo, reported by The Information, April 2026.
- 05 MLQ News, "Meta Caps Internal AI Token Spending After Costs Approach Billions in 2026," June 2026.
- 06 Praveen Neppalli Naga, reported via SmarterX / The Artificial Intelligence Show, June 2026.
- 07 SmarterX; Cockroach Labs, "The Bill Arrives: How to Manage Agentic AI Costs at Scale," June 2026.
- 08 Uber COO, reported via Business Insider, cited in Cockroach Labs, June 2026.
- 09 TechCrunch, "The token bill comes due," June 5, 2026; ADVISORI, June 2026.
- 10 TechCrunch, June 2026; Edgen, June 2026.
- 11 Optimum Partners, "AI Token Costs: Why Enterprise AI Bills Keep Rising in 2026."
- 12 Gartner, cited in Cockroach Labs and ADVISORI, March 2026.
- 13 Goldman Sachs, May 2026.
- 14 Goldman Sachs / SmarterX reporting, 2026.
- 15 Peter Sondergaard, 2021.AI AI.Watch, "The State of AI with Peter Sondergaard," 2025.
- 16 Ibid.