Point of View: AI Governance Is Broken. Here's How to Fix It.

Rick Hamilton, Jaswant Singh, and Naresh Nayar

January 20, 2026

I. The Problem: AI Governance as It Exists Today Is Failing

Organizations are deploying AI faster than they are learning to govern it, and the cracks are showing. With the last few years' explosion of generative AI solutions, what began as organizational experimentation has increasingly become operational dependence. We see this as AI now shapes underwriting decisions, clinical workflows, hiring pipelines, customer interactions, and strategic planning, across a variety of industries. Despite this, governance practices have not evolved at the same pace.

From our vantage point, many organizations still treat AI governance like traditional IT governance, with centralized control, technical oversight, and compliance checklists. Policies are drafted by senior committees, implemented by technical teams, and reviewed periodically for regulatory alignment. But this is not enough.

Our perspective is direct: this approach is fundamentally misaligned to how AI actually works, and how AI fails in operational scenarios.

AI systems, particularly agentic systems, are probabilistic and adaptive, and they are increasingly embedded across diverse business workflows. Their risks arise not only from code, but from context: how outputs are interpreted, which exceptions are ignored, where incentives distort behavior, and how small failures quietly accumulate all shape real-world outcomes. Traditional enterprise governance models assume predictability and linear cause-and-effect and, as a result, they systematically overlook the risks that matter most in AI-driven systems.

A further distinction between AI governance and prior IT governance lies in decision authority. Organizations must explicitly define which decisions AI may inform, which it may recommend, and which it may execute autonomously. These boundaries are not merely technical, but are organizational, ethical, and operational choices that evolve over time.

Effective AI governance must move at the same cadence as AI itself. Annual policy cycles and episodic reviews are misaligned with systems that learn, adapt, and act continuously. For agentic systems in particular, governance must extend into runtime operation, incorporating continuous supervision, real-time escalation signals, and the ability to pause, constrain, or override agents as conditions change.

Why Current Approaches Fall Short

Top-down governance is blind governance. Executive committees and centralized policy bodies operate far from where AI meets reality. They approve principles and frameworks, but they rarely see what matters most, including edge cases that only appear under real-world pressure; workarounds employees invent to “make the system work;” and those quiet failures that don't trigger alerts but erode trust over time. Eventually, when those problems surface to the top, the damage has often already been done.

Technical oversight alone misses the point. Accuracy, precision, drift detection, and model documentation are necessary, but not sufficient on their own. AI is not just a technical system; its successes and shortcomings have a strong behavioral element. Data scientists can tell you whether a model performs well on a test set. But they cannot always tell you whether an AI model's outputs are appropriate in a sensitive context; whether its users are over-trusting or under-trusting the AI model; or whether its use subtly shifts responsibility or accountability, and if so, in what way.

Thus, governance which focuses exclusively on technical control confuses correctness with business suitability. This suitability to accomplish business objectives is the foundational capability which must be kept top-of-mind.

Compliance-driven governance is reactive and shallow. Regulatory compliance is essential, but it typically represents the bare minimum, and not the requirements of a successful and advanced business operation. Laws lag AI's capabilities, so checklists reflect yesterday's risks, not tomorrow's needs. Organizations that equate compliance with governance tend to react after public failures, employee backlash, and in some cases, after regulators intervene. Thus, this approach conflates governance with damage control, not stewardship of business processes.

The Cost of Getting This Wrong

Regardless of the failure mechanism, when AI governance falls short, the consequences can be significant. These include reputational damage when AI misbehaves publicly; employee distrust that slows adoption and encourages “shadow AI;” regulatory exposure, particularly as global AI laws tighten; and most pervasively, wasted investment when promising AI initiatives stall or collapse.

AI governance setbacks are rarely catastrophic all at once. More often, they are cumulative, as small misalignments compound until the organization loses control of its own systems.

II. Our Point of View: The Three-Pillar Framework

Our core thesis is the belief that effective AI governance requires distributed accountability across three interconnected pillars:

1. First-line employee involvement in project selection, and in defining and monitoring proper AI behavior.

2. A cross-functional oversight committee that reviews KPIs, outcomes, and risks.
3. An independent audit function that red-teams AI use and challenges assumptions.

No single pillar is sufficient on its own. Together, these three functions form a system of checks and balances that reflects how AI operates inside organizations. In this context, governance defines decision rights, accountability, and escalation paths, while risk management implements controls and mitigations within the structure which governance establishes. For agentic AI, this system must also define bounded autonomy: clear thresholds for when agents may act independently, when human approval is required, and when authority must automatically revert to human control.

Why This Works

This framework deliberately combines ground truth from the people closest to AI use; strategic alignment from cross-functional leadership; and independent scrutiny from those empowered to question assumptions. It avoids the two most common governance failures: concentrating authority where visibility is weakest, and delegating responsibility without accountability.

This approach is not about slowing AI adoption; rather, it is about making AI adoption durable. Importantly, the parties entrusted with these multilayered responsibilities should each be action-minded and accountable; each pillar earns its place. Finally, this framework complements, rather than replaces, technical AI safety practices, and should not be treated as a substitute for pre-deployment evaluation, ensuring sufficient observability, or guaranteeing strong data security and privacy controls.

III. Pillar One: The Experts Are on the Front Line; Include Them

The people who use AI systems every day understand their real-world effects on business workflows better than anyone else. They see when outputs are helpful, when they are misleading, and when they quietly change staff behavior in unintended ways. Excluding these front-line employees from governance leaves you governing in the dark.

Staff Involvement in Project Selection

Front-line employees should have structured input into which AI projects move forward, and which should not. Although they may or may not see the broader financial picture, these employees know which business processes are broken or subpar; where automation would reduce risk or amplify it; and even in some cases, which decisions carry ethical, reputational, or safety sensitivities.

Practical mechanisms to solicit their input might include open proposal channels for AI ideas, structured prioritized inputs, and possible vetoes for high-risk use cases. This does not mean that first-line employees set AI strategy, but rather that such strategy is informed by operational reality as experienced by those working with the business processes and relevant tools.

Staff Role in Defining and Monitoring Proper AI Behavior

“Proper AI behavior” is difficult to succinctly define in a boardroom meeting. Rather, it emerges from employees' lived experiences with these tools, for better or worse.

Our view is that behavioral standards for AI must be co-created with the people who see its outputs every day. The AI criteria that employees can help define include domain-specific expectations (e.g., what “acceptable error” means in context), the boundaries for human override, and feedback loops to flag issues in real time. When employees see that their feedback leads to visible changes, trust increases, as does adoption readiness. Finally, if these front-line employees are not the users directly affected by the AI's decisions (e.g., the bank officer telling the customer that his loan application is denied), these employees should also act as the voice of the customer, providing valuable feedback on behalf of end users and affected parties.

Addressing the Counterarguments

The following might be heard as objections to involving employees. Each has a reasonable rebuttal. “Employees lack technical expertise”: perhaps, but employees are experts in outcomes, context, and impact, which the business needs. “This will slow AI progress”: employee inclusion is faster than recovering from a public failure or regulatory intervention. “They'll resist AI to protect their jobs”: involving staff in decision-making processes builds ownership of the AI models and their impact.

IV. Pillar Two: Strategic Oversight, with Diverse Organizational Perspectives

AI decisions rarely affect just one organizational function. Technical performance, legal exposure, workforce impact, customer experience, operational risk, and financial performance are deeply intertwined, and representatives from multiple areas must have a seat at the table.

Composition: Who Should Be Involved

Functional diversity, bringing different business perspectives and skill sets, must be designed in. At a minimum, the oversight committee should include technology leadership; legal / compliance; HR; operations; finance; customer-facing leadership; and ethics or risk (where applicable). In this role, seniority matters. Members should have authority to make binding decisions, not merely offer advice.

What the Committee Does

This body is not symbolic. With a regular cadence, e.g., monthly operational reviews and quarterly strategic assessments, its responsibilities include KPI and outcome review (performance and reliability, adoption and user satisfaction, risk indicators such as incidents, complaints, and bias flags, and business impact and ROI); pipeline oversight, evaluating proposed AI initiatives against values and risk tolerance; synthesis of front-line feedback, ensuring that ground truth informs strategy; and understanding end-user impact, qualifying and approving the end-user impact and ensuring that adequate means to gather reliable information exist.

Authority

An oversight committee without decision-making power is theater. This body must have a clear mandate to approve, pause, or stop initiatives. Further, it must report directly to executive leadership, and it must be accountable for outcomes, not just process adherence.

V. Pillar Three: Independent Scrutiny Through Red-Teaming

Governance without independent challenge often misses the mark. Every organization using AI needs someone whose job is to ask uncomfortable questions. For that reason, we strongly advocate red-teaming, or adversarial thinking, applied to governance. This panel's scope includes technical probing (stress-testing models, identifying edge cases, monitoring bias); process audits (verifying that governance procedures are followed, and are not just procedural lip service); outcome reviews (comparing real-world results to stated intentions); and cultural assessment (evaluating whether bad news travels upward, or is suppressed).

Questions That Matter

An effective audit function thinks holistically about the organization and AI's place within it. It should regularly ask about AI systems: what decisions is this AI actually making? How often is it wrong, and who bears the cost? How would we detect bias or drift over time? It asks about governance: is front-line feedback reaching decision-makers? Are incidents reported honestly or are they minimized? Do stated policies match actual practice? Are affected outside parties provided an appropriate user experience and are their voices heard? And it asks about culture: is there pressure to deploy faster than governance can keep up? Are people comfortable raising concerns, or are they afraid to?

Importantly, this function must report outside the teams it audits, to maintain independence. Best practices include having the audit function report directly to the C-suite or board. Ideally, the panel will entail a mix of internal expertise and external perspectives, allowing for diverse inputs on all elements of AI and its organizational integration. Finally, the audit committee's findings should be well-documented, with tracked remediation. This role must be recognized for its importance, and not simply procedural adherence.

VI. What This Framework Requires to Succeed

In our estimation, these three pillars are necessary, but not sufficient, for organizational AI success. Alongside these three pillars, organizations need executive accountability (an executive owner of AI governance should be named, to ensure that all elements are functioning properly); foundational principles (clarity on the values guiding AI decisions, so that all involved staff understand what is at stake and what is expected of the AI tools); technical visibility (at a minimum, each AI model must be supported by adequate documentation, version control, and explainability methods); incident

response capability (clear protocols for when something goes wrong, which may be independent of, or partially overlap, existing IT help desk protocols); and vendor and third-party coverage (governance must extend to AI you use but don't build, often including close coordination with vendors to ensure fit-for-purpose within your organization).

Other elements, including detailed regulatory mapping, technical safety methods, and change management mechanics are presumed, although detailing each is beyond the scope of this paper.

Finally, we must recognize that disagreements will occur. Put in place clear paths for resolving these disagreements: low-risk disagreements may be resolved at the team level; medium-risk conflicts may be resolved by the oversight committee; and high-risk or externally impactful conflicts may be escalated to an executive sponsor or board-level forum. Disagreement itself is not a failure of governance, but unresolved disagreement is. This framework assumes explicit decision rights, escalation paths, and accountability for overrides, particularly when speed, risk, and organizational incentives are in tension.

VII. Implications for Leaders

If you're just starting on your AI journey, we urge you to establish the framework for oversight first, as it's much easier to build within this existing framework, as opposed to bolting it on later. Put the mechanisms in place for front-line feedback, and build audit capability over time.

If you already have governance, map your current capabilities against the three pillars described here. Perhaps you have an oversight function, but have failed to adequately engage first-line employees, or perhaps you've done the latter but have no audit capability. Plan a path to seamlessly integrate the missing pieces to move yourself closer to this three-pillar approach.

If you're facing resistance, reframe the core question. Instead of “How do we prevent bad AI outcomes,” the question should be “How do we help teams deploy AI faster, safely, and with confidence?” Governance is a matter of enablement, and discussions can be tied back to risk and sustainability. Of course, demonstrating value on visible AI initiatives helps, so find a quick win or two to help underscore that AI successes and governance go hand in hand.

The bottom line is that organizations treating AI governance as overhead will be surpassed by those which treat governance as an enabling capability. The question before us is not whether to govern AI. It is whether to do the job well.

Conclusions

AI is too consequential to govern through good intentions or existing IT support capabilities alone. Since AI is not deterministic or fully specifiable in its outcomes, new approaches are needed to both maximize the upsides and mitigate risks. The organizations that will thrive are those that build

accountability into the fabric of their AI practices, grounding decisions in front-line reality, aligning them with strategic intent, and subjecting them to independent scrutiny.

This is our point of view. We welcome your input and the conversation to follow.

Originally published on Rick Writes AI (Substack), January 20, 2026.