# The Missing Control Loop: From the Boardroom to the Algorithm

### *Why AI governance must become a living organizational capability rather than a collection of promises*

**Maziyar Bahrami | July 22, 2026 | Research Essay**

In November 2023, the removal and rapid return of Sam Altman as chief executive of OpenAI transformed an internal dispute into a global governance experiment. The controversy was not only about leadership or personalities. It revealed a more consequential question: Can a private company govern a technology whose effects extend far beyond its investors, employees, and customers?

Roberto Tallarita used the OpenAI crisis to argue that artificial intelligence is testing the limits of corporate governance. OpenAI had deliberately departed from the conventional corporate model. Its nonprofit controlled structure was intended to protect the organization's social mission from direct investor pressure. Yet when the board attempted to exercise its unusual authority, pressure from employees, executives, investors, and the wider market rapidly weakened its position. The formal governance structure remained in place, but its practical authority proved fragile (Tallarita, 2023).

Tallarita's broader warning remains highly relevant. Traditional corporate governance was developed primarily to manage conflicts among shareholders, directors, and executives. It was not designed to govern technologies that may create large social, political, security, and systemic consequences. Moreover, separating a board from investor and executive influence does not necessarily make its decisions socially desirable. Independence creates space for judgment, but it does not guarantee competence, accountability, or wisdom (Tallarita, 2023).

The debate, however, has already moved beyond the question of whether corporate boards should oversee AI. Few serious observers now dispute that they should. The more difficult question is this:

> *How can responsibility expressed in the boardroom become effective control within the organization and, ultimately, within the AI system itself?*

This is where I see one of the most important research spaces in contemporary AI governance.

## The problem is not a shortage of principles

Organizations do not lack principles for responsible AI. Fairness, transparency, accountability, privacy, security, human oversight, and explainability appear in numerous corporate policies, regulatory documents, and professional guidelines.

The real difficulty begins when these principles must be translated into organizational decisions.

Who is allowed to approve an AI system? Who determines an acceptable level of risk? What information reaches the board? Who can delay or stop deployment? How are systems monitored after implementation? What happens when the model, data, regulation, or business context changes?

Papagiannidis, Mikalef, and Conboy (2025) make an important distinction between responsible AI principles and the governance practices required to implement them. Their research organizes responsible AI governance into three types of practices: structural, procedural, and relational. Structural practices allocate authority and responsibility. Procedural practices establish assessment, monitoring, and decision processes.

Relational practices connect the technical, managerial, legal, and social actors involved in AI development and use.

Their review also reaches a significant conclusion. Although organizations increasingly adopt responsible AI principles, our understanding of how those principles are operationalized during the design, execution, monitoring, and evaluation of AI systems remains limited and fragmented (Papagiannidis et al., 2025).

This distinction is fundamental. Publishing principles is not the same as governing AI. Creating an ethics committee is not the same as controlling AI. Mentioning AI risk in an annual report does not demonstrate that an organization can identify, escalate, and correct that risk.

Governance should therefore be assessed not only through the sophistication of its language, but through the organization's ability to act.

## When governance looks stronger than it is

One of the most revealing studies for understanding this problem does not examine AI directly. Lowry, Vance, and Vance (2025) investigate corporate boards' oversight of cybersecurity, another technical, rapidly changing, and strategically important organizational risk.

Their findings are striking. Directors without sufficient cybersecurity expertise may genuinely attempt to provide diligent oversight. They may receive reports, ask questions, attend committee meetings, and follow accepted governance routines. Yet their supervision can remain largely symbolic because they cannot independently evaluate the information provided by management, recognize weak answers, or determine whether apparently reassuring controls are genuinely effective.

The visible governance activities may look similar, but their substance is different. Expert and nonexpert directors can perform the same formal activities while exercising very different levels of meaningful oversight (Lowry et al., 2025).

This finding has direct implications for AI governance.

A board may discuss AI at every meeting and still be unable to challenge management's assumptions. A company may establish an AI committee that lacks access to relevant information, technical competence, executive support, or decision authority. An organization may produce extensive documentation without possessing a credible mechanism for suspending a dangerous system.

Recent evidence from corporate disclosure suggests that this concern is justified. Marin et al. (2025) analyzed more than 30,000 regulatory filings from over 7,000 publicly listed companies in the United States. The proportion of companies discussing AI risk increased sharply between 2020 and 2024. By 2024, approximately 43 percent mentioned AI in the risk factors section of their annual filings. Nevertheless, many disclosures remained generic or included little information about concrete mitigation measures (Marin et al., 2025).

The central governance challenge is therefore not simply whether governance exists. It is whether governance is symbolic or substantive.

Symbolic AI governance communicates that an organization is attentive, modern, responsible, and compliant. Substantive AI governance changes decision rights, resource allocation, system design, monitoring practices, escalation pathways, and organizational behaviour.

The two can coexist. A company can adopt meaningful controls while also using governance language to protect its reputation or legitimacy. The research challenge is to identify when formal governance mechanisms represent a real organizational capability and when they primarily provide reassurance.

## An AI committee cannot govern alone

Organizations have begun creating algorithm review boards, ethics committees, and related internal bodies. These mechanisms can be valuable, but their existence is not sufficient.

Hadley, Blatecky, and Comfort (2025) interviewed 17 technical contributors working across government, industry, nonprofit, and academic organizations. Their study provides rare empirical evidence about algorithm review boards in practice. The review boards differed greatly in membership, authority, scope, and procedures. Some examined every model, while others reviewed only selected documentation or high risk applications.

Two conditions appeared particularly important. Review boards needed to be integrated with existing organizational processes, and they needed genuine support from senior leadership. Financial tension was also persistent because responsible AI practices require time, expertise, and resources that may conflict with short term commercial pressures (Hadley et al., 2025).

This suggests that an AI committee should not be treated as an isolated solution. It must be connected to enterprise risk management, cybersecurity, data governance, internal audit, legal functions, system development, and strategic decision making.

Most importantly, it must possess sufficient authority to influence whether an AI system is developed, deployed, restricted, redesigned, or suspended.

A committee that can discuss risk but cannot change a decision is not a governance mechanism. It is an advisory forum.

## Expertise matters, but expertise alone is insufficient

A natural response to these concerns is to demand more AI experts on corporate boards. Greater expertise is necessary, but it is not a complete solution.

Tallarita (2023) introduces the concept of cognitive distance to explain this problem. Boards need directors with different experiences, forms of knowledge, and interpretations of risk. Too little cognitive distance can produce groupthink and intellectual conformity. Too much distance can prevent communication, mutual understanding, and collective judgment.

An effective board may therefore need more than one technically knowledgeable director. It may require a combination of technical, strategic, legal, ethical, cybersecurity, and organizational expertise. It also needs processes that allow these different forms of knowledge to interact productively.

The relevant question is not merely whether expertise exists somewhere in the boardroom. The question is whether the organization can convert that expertise into decisions, controls, and accountability.

A related issue arises when boards use AI to support their own decisions. Kourabas and Tsang (2025) argue that AI may help boards address persistent limitations involving time, information, and board composition. At the same time, reliance on AI may weaken accountability if directors defer excessively to automated recommendations.

They therefore distinguish between selecting the right humans to oversee AI assisted decisions and defining the right process through which those humans exercise judgment. Directors may use AI, but they remain responsible for the resulting decisions (Kourabas & Tsang, 2025).

Human oversight should consequently not be treated as a checkbox. Its quality depends on who performs it, what information they receive, what authority they possess, and whether they can recognize when intervention is necessary.

## From governance structure to governance capability

My central argument is that AI governance should be understood as an adaptive organizational capability, not simply as a policy, committee, or reporting structure.

Such a capability allows an organization to perform six connected activities.

First, it senses change. The organization identifies developments in AI technology, regulation, organizational use, threats, and stakeholder expectations.

Second, it maps exposure. It understands where AI is used, what data and infrastructure each system depends on, who may be affected, and how risk can move across organizational and technological boundaries.

Third, it decides. It establishes risk tolerances, responsibilities, decision rights, and the conditions under which development or deployment may proceed.

Fourth, it controls. It translates these decisions into testing requirements, access restrictions, monitoring systems, documentation, human approval, limits on autonomy, and technical safeguards.

Fifth, it responds. It detects deviations, escalates concerns, investigates incidents, restricts systems, and corrects failures.

Finally, it learns. It uses incidents, near misses, audits, regulatory changes, and operational evidence to redesign both the AI system and its governance arrangements.

The NIST Artificial Intelligence Risk Management Framework provides an important foundation for this perspective. Its four functions, Govern, Map, Measure, and Manage, are intended to support continuous risk management throughout the AI lifecycle rather than a single assessment before deployment (Tabassi, 2023).

An organizational capability perspective goes one step further. It asks whether information and action actually flow between these functions.

A board may define a risk tolerance that never reaches system developers. A technical team may identify model drift that is never communicated to senior management. An audit may reveal a weakness without changing budgets, responsibilities, or system permissions.

In each case, the individual components of governance may exist, but the overall control loop is broken.

The central problem is therefore the distance between boardroom intention and system behaviour.

## Agentic AI makes the gap more dangerous

This problem becomes more urgent as organizations adopt agentic AI systems that can plan tasks, select tools, access enterprise data, and execute sequences of actions with limited direct supervision.

With conventional predictive AI, governance often focuses on how an output is produced and how a human uses it. With agentic AI, governance must also determine what the system is permitted to do.

Which databases may it access? Which tools may it use? Can it modify organizational records? How long can it retain information? When must it request human approval? Who can remove its authority?

Dux et al. (2026) show that the governance of agentic AI is implemented through concrete architectural and organizational arrangements. These include decisions about tool access, data boundaries, memory, human approval, traceability, learning, and the gradual expansion of autonomy. Governance is not an external

compliance layer placed around a completed system. It is part of the system's architecture and its process of organizational implementation (Dux et al., 2026).

This changes the practical meaning of corporate governance. Board responsibility cannot end with the approval of an AI policy or the receipt of a quarterly report. The organization's risk appetite must eventually be translated into permissions, monitoring thresholds, approval requirements, escalation rules, and intervention mechanisms.

Governance must travel from the boardroom into the architecture. Evidence from the architecture must also travel back to the boardroom.

## The research space I find most promising

For Information Systems and management research, the next step should not be the creation of another general list of responsible AI principles. The more valuable task is to explain and measure what AI governance actually does inside organizations.

First, we need better measures of substantive AI governance capability. Researchers can combine governance disclosures, board biographies, committee responsibilities, internal policies, incident reports, litigation, system documentation, and technical assurance practices. Natural language processing may help distinguish organization specific mechanisms from generic commitments.

Second, we should investigate how the cognitive configuration of a board influences governance effectiveness. The question is not simply whether one director has AI expertise. It is how technical, strategic, cybersecurity, legal, and ethical knowledge are distributed and integrated. Both cognitive uniformity and excessive cognitive distance may weaken meaningful oversight.

Third, AI governance should be connected more directly with cybersecurity governance. AI systems depend on data security, access management, third party controls, incident response, infrastructure resilience, and protection against adversarial behaviour. Firms with mature cybersecurity governance may be better prepared to govern AI. However, they may also develop false confidence if they assume that AI risk is only another form of information security risk.

Finally, governance research must move beyond adoption and disclosure to examine outcomes. Do AI committees prevent or detect failures? Does relevant board expertise improve escalation and correction? Do governance mechanisms increase trust without unnecessarily obstructing useful innovation? Can firms demonstrate that their controls function when commercial pressure is high?

These questions are difficult because effective governance often prevents events that never become visible. Yet without serious attempts to evaluate outcomes, governance frameworks may become increasingly sophisticated descriptions of activities whose effects remain unknown.

## Governance must become a living system

Tallarita concludes that corporate governance cannot adequately manage catastrophic AI risk and that private governance cannot replace public authority. Firms cannot be expected to resolve alone every conflict among technological competition, economic incentives, and social welfare (Tallarita, 2023).

Still, a large and important organizational space exists between unrestricted technological development and direct government control. Companies decide which systems to build, what data those systems can access, how much autonomy they receive, how failures are detected, and when intervention becomes mandatory.

The next generation of AI governance must therefore be more than a statement of values. It must become a living system of authority, information, technical control, accountability, and learning.

The defining governance question is no longer simply:

> *Who is responsible for AI?*

It is:

> *Can responsibility move quickly and reliably enough from the boardroom, through the organization, and into the architecture to influence what AI systems actually do?*

Until organizations can answer that question with evidence rather than promises, the missing control loop between the boardroom and the algorithm will remain one of the most consequential governance gaps of the AI era.

## References


Dux, N., Alaimo, C., Roussiere, P., & Mishra, A. K. (2026). Governance by design: Architecting agentic AI for organizational learning and scalable autonomy. arXiv preprint arXiv:2605.20210. doi: 10.48550/arXiv.2605.20210.

Hadley, E., Blatecky, A., & Comfort, M. (2025). Investigating algorithm review boards for organizational responsible artificial intelligence governance. AI and Ethics, 5, 2485–2495. doi: 10.1007/s43681-024-00574-8.

Kourabas, S., & Tsang, C. Y. (2025). The board monitoring function: Artificial intelligence in the era of heightened accountability. European Corporate Governance Institute Law Working Paper No. 856/2025. doi: 10.2139/ssrn.5200732.

Lowry, M. R., Vance, A., & Vance, M. D. (2025). Inexpert supervision: Field evidence on boards' oversight of cybersecurity. Management Science, 72(2), 783–804. doi: 10.1287/mnsc.2023.04147.

Marin, L. G. U. B., Rijsbosch, B., Spanakis, G., & Kollnig, K. (2025). Are companies taking AI risks seriously? A systematic analysis of companies' AI risk disclosures in SEC 10-K forms. Proceedings of the SoGood Workshop, ECML PKDD 2025. arXiv:2508.19313.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. The Journal of Strategic Information Systems, 34(2), Article 101885. doi: 10.1016/j.jsis.2024.101885.

Tabassi, E. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology, NIST AI 100-1. doi: 10.6028/NIST.AI.100-1.

Tallarita, R. (2023, December 5). AI is testing the limits of corporate governance. Harvard Business Review.