

# Who Is Accountable When AI Makes the Decision?

---

**AI can recommend, decide, and increasingly act. But authority can move to a machine faster than accountability can move through an organization. The executive problem is not whether AI can own a task; it is whether the enterprise can still identify who owns the outcome when the system acts.**

Executive Operations Insights | Saša Cibula

Imagine a regional operations network using an AI planning agent to allocate scarce inventory during a supply disruption. The agent evaluates customer demand, margin, contract commitments, service risk, transportation capacity, and inventory position. It then reallocates product across six distribution centers and changes several customer commitments before a planner intervenes.

The economic result is defensible. Total service improves, premium freight falls, and the network protects more revenue than the prior manual process usually achieved. But one strategically important customer receives less product than its commercial team expected, creating an escalation that reaches the executive team.

Who owns that decision? The commercial leader whose priorities were encoded into the objective? The operations leader who deployed the process? The technology team that built the agent? The risk function that approved the controls? The planner who did not intervene? Or the model itself?

The last answer is the only one that is clearly wrong. **AI can receive decision authority. It cannot absorb organizational accountability.** If an enterprise cannot identify who remains accountable when an AI-enabled process produces a consequential outcome, the governance design is incomplete.

## Accountability Cannot Be Delegated to a Model

NIST's AI Risk Management Framework provides a strong starting point because it treats governance as a cross-cutting organizational responsibility rather than a technical control added after deployment. Within the Govern function, NIST calls for clear roles and lines of communication, explicit accountability structures, defined human-AI oversight, and executive responsibility for decisions about AI risk and deployment.

That framing matters. When a human manager makes a decision, accountability is usually inferred from role, authority, policy, and organizational hierarchy. As AI becomes more autonomous, those assumptions weaken. The system may synthesize information from multiple functions, apply objectives set by different leaders, use models supplied by vendors, and act inside workflows that span several business units. The decision is produced by a system, but accountability still has to terminate in the organization.

The governance requirement therefore becomes more explicit as autonomy increases, not less. Leaders need to know who owns the business outcome, who has authority over the decision class, who is responsible for the system's operation, who monitors performance and risk, and who can suspend or override the system when its boundaries are breached.

## NIST AI RMF: ACCOUNTABILITY IS GOVERNED ACROSS THE AI LIFECYCLE

Govern is cross-cutting: roles, executive responsibility, oversight, and escalation must remain explicit as AI moves from recommendation to action.

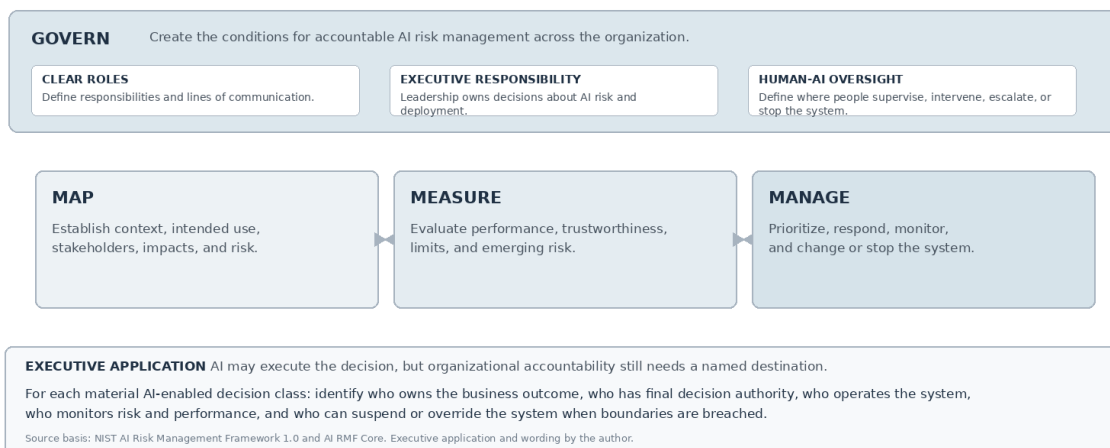


Figure 1. NIST AI RMF governance applied to organizational accountability for AI-enabled decisions.

## Accountability Is Larger Than Model Ownership

One of the easiest mistakes is to assign accountability to the team that owns the AI model. That team may be responsible for technical performance, data pipelines, monitoring, access controls, or deployment. It rarely owns the business objective that the AI is optimizing.

An operations planning agent, for example, may be technically owned by a data or technology organization while optimizing decisions involving inventory, labor, service, customer priority, and transportation cost. Those are business choices. If the model is instructed to maximize contribution margin subject to a service floor, someone in the business has made a policy decision about the relative importance of margin and service. The algorithm expresses that policy at scale; it does not create the organization's accountability for it.

This distinction becomes critical during incidents. A technically healthy system can make an organizationally unacceptable decision because the objective, constraints, or decision rights were poorly defined. Conversely, an operationally desirable objective can still fail if the model, data, or controls are defective. Accountability therefore has multiple components, but that does not mean accountability should become diffuse.

## RAPID Clarifies the Human Decision Architecture

Bain's RAPID framework remains useful because it separates the roles required to make and execute a decision: Recommend, Agree, Perform, Input, and Decide. Its most important discipline is that a critical decision should have a clear "D" - a person or group with final authority to commit the organization to action.

AI changes how those roles can be performed. A model can gather inputs, develop a recommendation, and in some cases perform the resulting action. But the presence of automation does not eliminate the need to identify who has the organizational authority to define the decision rules and accept the resulting risk.

For a material AI-enabled decision, executives should be able to answer several questions without convening a crisis meeting. Who has the final authority over this class of decision? Who determines the objective the AI is optimizing? Who is responsible for validating that the system remains within approved boundaries? Who performs the operational response when the system escalates? Who can stop it?

Those responsibilities may sit with different people. The important point is that they are explicit. "The AI made the decision" cannot become a new form of organizational ambiguity.

## The Human-in-the-Loop Can Create False Accountability

Many organizations respond to AI risk by inserting human approval. That can be appropriate, especially for high-consequence or ambiguous decisions. But the presence of a person in the workflow does not automatically create meaningful accountability.

If a planner receives hundreds of AI recommendations each day, has seconds to review each one, rarely sees the underlying reasoning, and is measured primarily on throughput, the formal approval may be little more than a click. The organization can then claim that a human made the final decision even though the system effectively determined the outcome.

This produces the worst of both models: machine-scale decisions with human-scale latency, plus ambiguous accountability when something goes wrong. The human becomes responsible on paper without possessing enough time, information, or authority to exercise real judgment.

The better design is to put accountability around the system rather than pretending every transaction requires human authorship. Leaders define the objective, permitted actions, decision thresholds, escalation triggers, monitoring requirements, and conditions that force authority back to a person. Low-risk decisions can then execute automatically while human attention is concentrated where judgment and accountability materially change the outcome.

## A Practical Example: Allocation During a Shortage

Return to the six-site network facing a product shortage. The AI agent has been authorized to rebalance inventory automatically across facilities and customer orders. Before deployment, leadership has established a hierarchy of service commitments, contractual protections, margin objectives, and inventory floors. Routine rebalancing decisions happen without human approval.

During a severe shortage, however, the system begins trading service between major customers. The optimization is no longer merely operational. It is making choices with commercial, reputational, and potentially contractual consequences. At that point, the decision class should cross an authority boundary.

Under a well-designed accountability structure, the AI can continue to map options, quantify consequences, and recommend the economically strongest response. The operations organization may own the process and perform the allocation. Commercial leaders provide input on customer implications. Legal or risk may have an Agree role where contractual obligations require it. But one executive still has the D for the exceptional allocation policy and accepts the business tradeoff.

The machine may generate the decision faster than any team could. It may even execute it once the policy boundary is established. Yet the organization has not outsourced accountability. It has clarified where accountability changes as the consequence of the decision changes.

## Accountability Is an Economic Control, Not Just a Governance Requirement

This is not only about compliance or blame. Clear accountability improves operating speed. When decision rights are ambiguous, organizations compensate with committees, manual approvals, duplicated reviews, and escalations. Every additional checkpoint protects against uncertainty but also adds cost and latency.

The opposite failure is equally expensive. When autonomous systems operate without explicit ownership, organizations discover problems late, spend more time reconstructing why decisions occurred, and often respond by imposing broad restrictions that slow every AI use case. Weak accountability can therefore reduce both trust and return on the technology investment.

Recent enterprise research reflects this relationship. McKinsey's 2026 AI trust research found that organizations with explicit ownership for responsible AI reported materially higher maturity than those without clear accountability. BCG's work on enterprise AI control similarly emphasizes centralized visibility, common policy enforcement, named ownership, and the ability to intervene or shut down agents as autonomy scales.

Good governance should therefore make safe action faster, not simply make unsafe action harder. Clear accountability is one of the mechanisms that makes greater autonomy possible because leaders know who sets the boundaries, who watches the system, and who owns the consequences.

## The Counterargument: AI Decisions Are Inherently Shared

There is a legitimate objection to forcing single-point accountability onto complex AI systems. Outcomes may depend on data owners, model developers, vendors, process owners, users, risk teams, and executives. No individual can realistically control every contributor. Treating one person as solely responsible can oversimplify a genuinely distributed system.

That objection is correct if accountability is confused with sole responsibility. Complex systems require shared responsibilities. NIST explicitly recognizes organizational teams, partners, and multiple risk-management roles. ISO/IEC 42001 likewise treats responsible AI as a management system involving policies, processes, responsibilities, monitoring, and continual improvement across the organization.

But shared responsibility does not justify ambiguous accountability. A commercial executive does not need to own model engineering to be accountable for the customer-allocation policy. A technology leader can be accountable for platform controls without owning the business outcome. An operations leader can own the performance of the AI-enabled process while independent assurance remains with risk or audit.

The goal is not to find one person to blame. It is to ensure that every material decision has a clear organizational destination for authority, risk acceptance, execution, and oversight.

## The More Autonomous the System, the More Explicit Accountability Must Become

As AI moves from analysis to action, organizations will increasingly separate the person who defines the policy from the system that executes it. That can create enormous productivity and speed advantages. It can also create an accountability gap if executives assume that automated execution somehow reduces the need for human ownership.

It does the opposite.

The more frequently AI acts without transaction-level approval, the more important it becomes to define the decision boundaries, identify the owner of the business outcome, assign final authority for the decision class, establish independent monitoring, and preserve a clear right to intervene.

The question after a consequential AI decision should never be, “Who was supposed to be watching the model?” It should already be known who owns the decision, what authority the system was given, why those boundaries were chosen, and who accepted the risk.

AI may make the decision. **The enterprise still owns the consequence.**

## Selected Sources

**National Institute of Standards and Technology.** *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. January 2023. [Source](#)

**National Institute of Standards and Technology.** *AI RMF Core - Govern, Map, Measure, Manage*. [Source](#)

**ISO.** *ISO/IEC 42001:2023 - Artificial Intelligence Management Systems*. December 2023. [Source](#)

**Bain & Company.** *The Five Steps to Better Decisions*. [Source](#)

**McKinsey & Company.** *State of AI Trust in 2026: Shifting to the Agentic Era*. March 2026. [Source](#)

**Boston Consulting Group.** *Enterprise AI Control Plane: The CIO's Guide to Governing and Accelerating AI Agents*. August 2026. [Source](#)