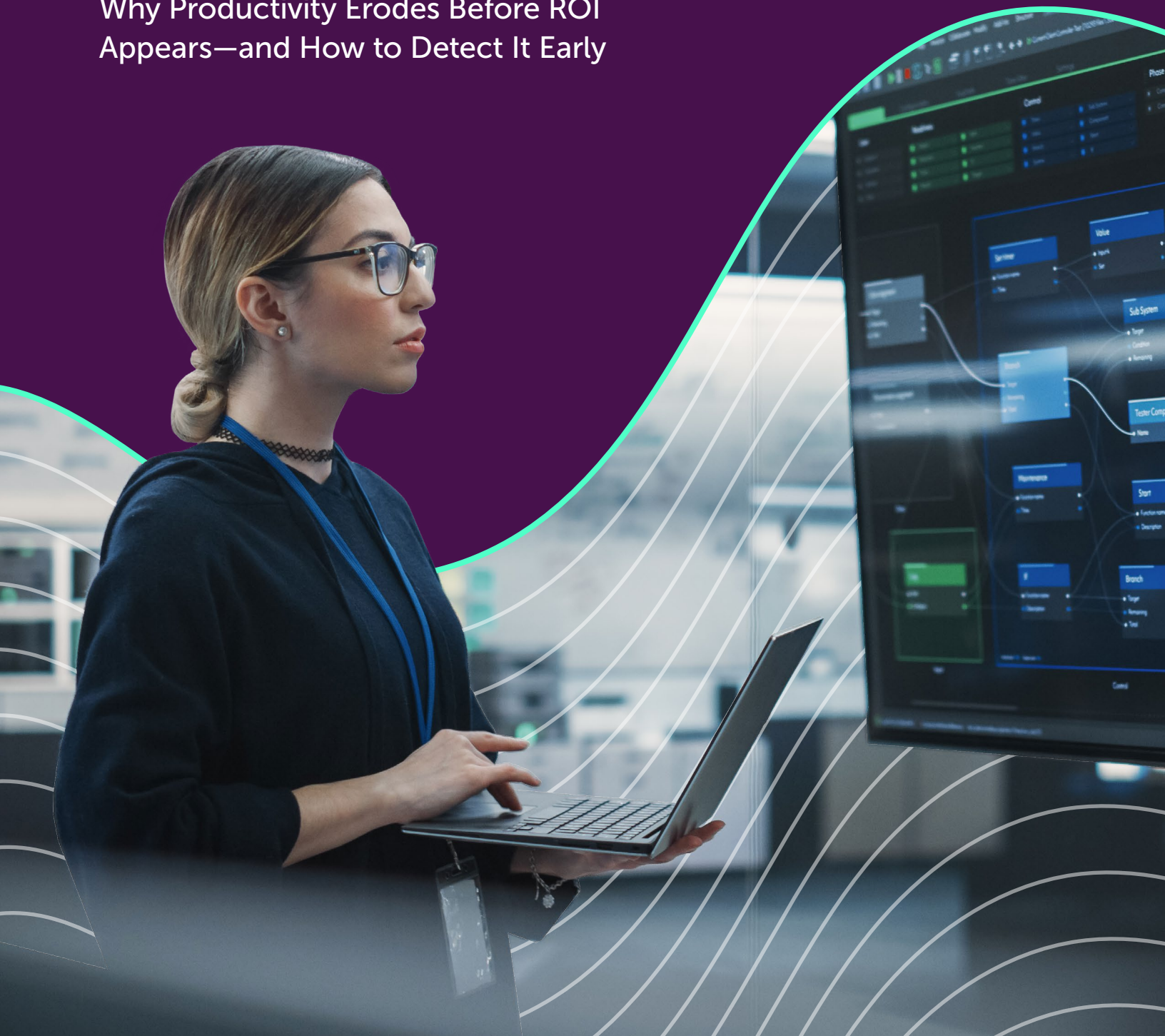
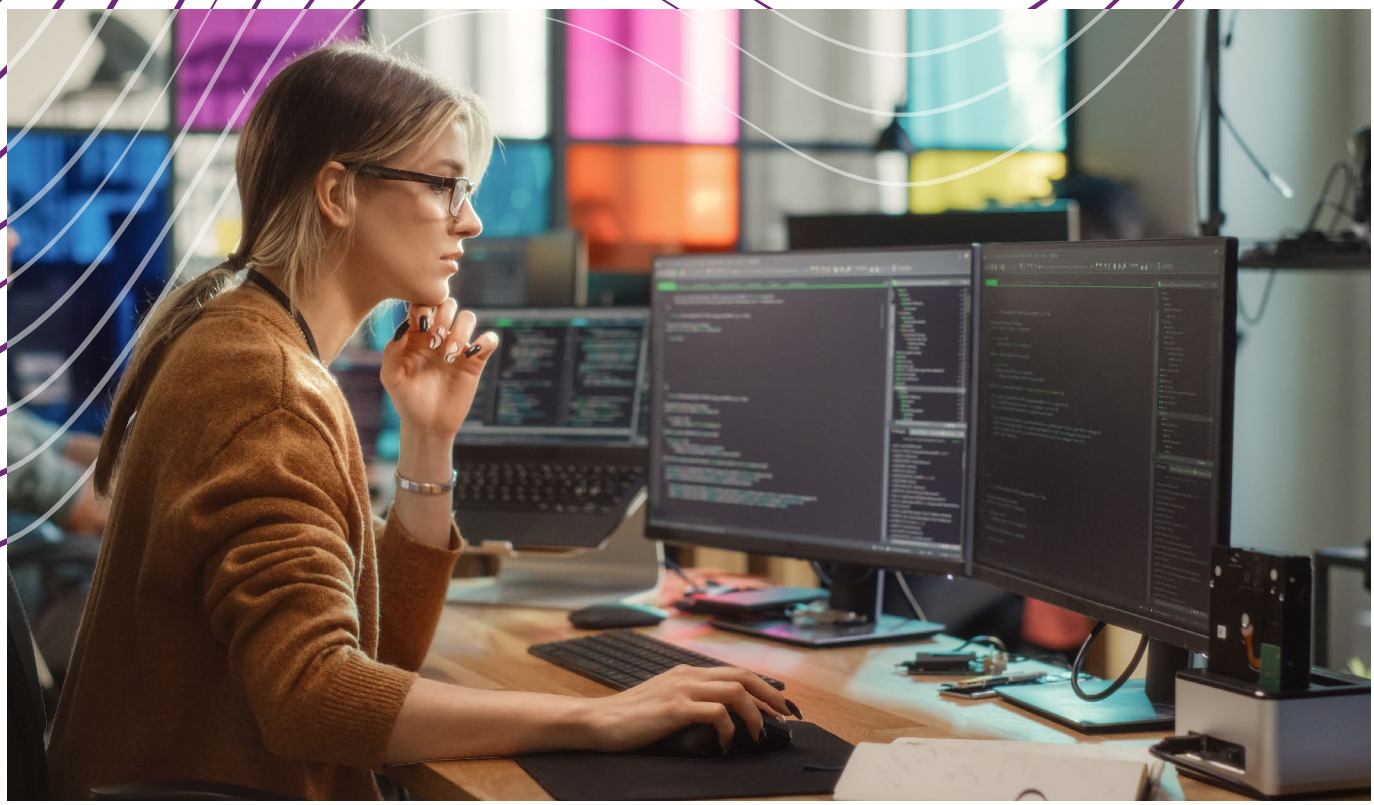
seramount
Talent Solutions for Better Workplaces

# The AI Execution Risk Index

## Why Productivity Erodes Before ROI Appears—and How to Detect It Early



# Contents



# **The Illusion:** When AI Gains Do Not Compound

Enterprise investment in generative AI has accelerated rapidly over the past three years. Global spending on AI systems is projected to exceed $2 trillion this year, and more than two-thirds of large organizations now report regular use of generative AI in at least one business function (Gartner, 2026; McKinsey, 2025). As investment and adoption accelerate at unprecedented speed, AI is frequently cited in earnings calls and board discussions as a primary lever for improving productivity, efficiency, and margin expansion.

At the individual task level, this optimism is supported by empirical evidence in targeted use cases. In a study of more than 5,000 customer support agents, AI helped teams resolve issues faster—improving performance by 14 percent overall and by more than 30 percent for lower-performing employees—without hurting customer satisfaction (Brynjolfsson et al., 2024). Furthermore, studies in software and human-computer interaction show that generative AI does more than accelerate tasks; it redistributes cognitive effort, reshapes managerial oversight, and alters how synthesis and evaluation occur inside organizations (Hoffmann et al., 2024; Lee et al., 2025). For execution and operational leaders, these gains translate into a compelling promise: higher throughput without increased headcount, faster cycle times, and potential reductions in cost-to-serve.

Yet aggregate productivity data tells a more complicated story. Economy-wide performance growth has not reflected dramatic acceleration, despite rapid enterprise deployment (Humlum and Vestergaard, 2025). Studies indicate that while adoption is widespread, only a small minority of organizations report achieving enterprise-wide AI maturity or consistent, measurable business impact (BCG, 2024). Many initiatives remain confined to pilots, and failure rates for digital and AI transformation programs remain high across industries (MIT, 2025).

This divergence between individual gains and organizational outcomes is not new. It echoes what economists once described as the IT productivity paradox: Technology appears everywhere except in productivity statistics (Brynjolfsson, Rock, and Syverson, 2021). What makes the current moment distinct, however, is that generative AI does not merely accelerate execution; rather, it operates inside cognitive processes that were once exclusively human. It drafts, synthesizes, analyzes, and proposes decisions, participating directly in how judgment is formed. **Rather than wholesale job elimination, the dominant pattern recent labor market analysts are tracking is task redistribution: AI shifts which activities humans perform and how those activities are sequenced** (Anthropic, 2026). This redistribution complicates how productivity gains coalesce at the enterprise level.

In short, generative AI collapses the cost of production, expands the volume of potential output, and introduces a new set of enterprise tensions regarding how to organize human-machine capability. When drafting, coding, and analysis become dramatically cheaper, activity increases almost immediately. But the capacity to evaluate, prioritize, contextualize, and own that work has not risen at the same pace. Without explicit redesign of decision rights, skills architecture, and accountability norms, organizations risk amplifying activity faster than they strengthen execution stability (Deloitte, 2026).

These findings do not indicate that AI is ineffective; rather, they indicate that value realization is conditional. Across studies and industries, researchers consistently point to the following two tensions inhibiting true ROI:

### 1. Organizational redesign lags behind technological deployment.

### 2. Measurement systems emphasize usage rather than system-level stability.

In other words, the mechanism limiting AI ROI is not technological fluency. It is organizational absorption and enterprise evolution. AI improves tasks faster than organizations improve systems.

We refer to this structural instability as **AI Execution Risk**—the risk that AI deployment destabilizes operating systems before it strengthens them. With AI already delivering measurable gains at the task level, this paper introduces a structured index to help executive teams identify where those gains fail to translate into enterprise performance—and how to stabilize execution before scaling.

# **The Imbalance:** Where AI Creates Structural Strain

To understand why AI ROI frequently stalls, it is necessary to move beyond adoption metrics and examine what changes occur inside operating systems when AI is introduced. Across industries, research shows that three structural imbalances tend to emerge during early and mid-phase AI deployment.

## 1 Acceleration Outpaces Workflow Redesign

Research consistently shows that most organizations deploy AI to optimize existing workflows rather than to redesign them (Korst et al., 2025). Legacy processes remain intact. Approval chains persist. Review cycles are preserved as risk controls.

In this configuration, AI is often layered onto current workflows instead of replacing them. The result is not immediate simplification. It is volume expansion within preexisting complexity. This dynamic is sometimes described as "workflow debt": accumulated procedural complexity that predates AI and ultimately constrains its impact (Bain & Company, 2026). **When AI accelerates production within indebted systems, it amplifies inefficiency rather than eliminating it.**

The operational consequence is subtle—at first. Employees produce more drafts, more analyses, and more iterations. While time per task decreases, the number of handoffs and review cycles often increases. The system moves faster—but not better. Acceleration without workflow redesign converts to margin pressure.

## 2 Output Outpaces Judgment

The appeal of AI is rooted in its dramatic reduction in production cost. But AI does not reduce the cost of evaluation. The constraint shifts from producing work to evaluating it.

When output expands without a corresponding redesign of review thresholds, decision rights, and performance standards, organizations experience what can be thought of as "review inflation": more deliverables without proportional improvement in outcome quality.

In operational terms, the bottleneck migrates, especially when workflow redesign fails to take shape. Managers spend more time adjudicating AI-assisted outputs followed by an increase in escalations, and ambiguity around ownership expands. **Without explicit recalibration of judgment systems—including clearer decision and accountability protocols—work quality declines.**

Acceleration at the task level can therefore expose a gap in judgment capacity. What appears as productivity locally can introduce volatility system-wide.

## 3 Workflow Complexity Outpaces Coordination

If the first two imbalances describe pressure inside workflows and decision systems, this third dynamic operates at the enterprise level. Enterprise AI rarely enters through a single coordinated rollout. Adoption spreads function by function, often originating within business units rather than centralized transformation programs (McKinsey, 2025). The result is not just layered workflows—it is a breakdown in how work connects across the enterprise.

As AI tools spread across teams, more work must move across functions. Outputs created in one area need to be reviewed, adjusted, or approved in another. But coordination norms rarely evolve at the same pace. The result is friction. Teams use different tools, data must be reconciled, and more time is spent aligning decisions across functions. Organizations often assume they are ready to expand a tool's use based on adoption (*HBR*, 2026). In practice, their ability to coordinate that work has not caught up, because cross-functional initiatives require more synchronization (Deloitte, 2023). What looks like acceleration at the team level can translate into slower execution at the enterprise level. Time-to-value is extended as integration friction absorbs early productivity gains.

For operational leaders, this becomes a throughput constraint. **When coordination capacity does not proportionately increase alongside deployment, acceleration at the edge produces strain at the center.** What appears as local efficiency translates into enterprise-wide integration overhead.

# **The Risk:** Execution Volatility and Cost-to-Serve

Individually, the imbalances described above appear manageable. A few additional review cycles. More iterations. Slightly heavier coordination. Each signal, in isolation, looks like the normal friction of innovation. Collectively, however, they produce a more consequential pattern: execution volatility. Execution volatility emerges when acceleration outpaces structural adjustment. When usage of AI is scaled inside these conditions, existing weaknesses in workflow, decision-making, and coordination become more pronounced.

This dynamic is not new. It echoes long-standing automation research on human-computer collaboration (Bainbridge, 1983), which showed that as routine tasks are automated, the remaining human work becomes more complex, less predictable, and more cognitively demanding. What is different now is the scale and scope. Contemporary studies of AI-enabled work show that this shift is occurring across a far broader set of roles, increasing oversight burden and creating greater ambiguity around accountability (Hoffmann et al., 2024; EPRS, 2025). As routine drafting, summarizing, and analytical tasks are accelerated, the remaining human work does not simply shrink—it changes. Workers are left handling exceptions, complex judgment calls, cross-functional trade-offs, and situations that don't fit a clear rule.

In other words, as systems accelerate, the decisions that remain human become fewer in number but greater in complexity. As recent Seramount research shows, those decisions require interpretation rather than execution, adjudication rather than production. In AI-enabled environments, the easier it becomes to generate options, the more demanding it becomes to evaluate, prioritize, and own their consequences. The cognitive burden does not disappear with automation; it is concentrated, ultimately affecting cost-to-serve.

**Cost-to-serve increases when:**

- Review layers expand to mitigate uncertainty.
- Rework cycles lengthen as AI-generated output requires refinement.
- Meeting load rises to reconcile cross-functional inconsistencies.
- Managerial bandwidth is consumed by evaluation and integration tasks rather than value creation.

None of these shifts immediately register on traditional productivity dashboards. Adoption metrics remain positive. Activity levels remain high. Output per employee may even increase. But financial indicators eventually follow operational strain.

When AI is scaled without assessing system absorption capacity, acceleration can magnify preexisting fragilities in workflow design, decision architecture, and coordination norms. **What begins as productivity gain can evolve into execution instability—expressed through volatility, variance, and rising cost-to-serve.** For executive teams, the core issue comes down to the durability of the operating model. Increased output creates advantage only if the system can absorb it without eroding reliability or margin.

# **The AI Execution Risk Index:** Assessing for Gaps

By the time organizations begin evaluating AI ROI, adoption is already embedded. Tools are integrated into workflows. Expectations have shifted. Scaling decisions are underway.

At that stage, leadership responsibility shifts from experimentation to structural readiness.

Enterprise value depends on whether the operating model—workflows, decision rights, coordination systems, and capability pathways—evolves alongside AI deployment. Adoption metrics offer limited insight into that question. What determines durable performance is execution stability.

The AI Execution Risk Index provides a structured lens for examining whether acceleration is outpacing adaptation. It identifies five organizational fault lines where instability typically emerges before financial impact becomes visible.

## 1. Workflow Absorption

*Can the system eliminate work—or is it only accelerating it?*

AI frequently increases production before legacy work is removed. When workflow simplification falls behind deployment, automation is layered onto complexity rather than replacing it. At scale, this does not simply increase activity; it compounds structural cost. The issue is not speed. It is structural redundancy.

### How Instability Surfaces

- Teams generate more drafts and iterations, but approval structures remain unchanged.
- Review timelines are compressed while total process steps increase.
- AI outputs coexist with parallel manual verification and legacy reporting safeguards.

In these environments, efficiency gains are consumed by static process overhead.

### Structural Examination

- What work has been formally eliminated since AI deployment began?
- Have approval thresholds been reset—or merely accelerated?
- Are safety nets masking unresolved workflow debt?

### Leadership Imperative

Remove friction before you multiply output. Acceleration without structural simplification converts efficiency into cost pressure.

## 2. Decision Rights

*Has accountability evolved alongside machine participation?*

As AI contributes to drafting, analysis, and recommendation generation, the value of critical thinking and judgment increases, including when to use AI and when to not. Yet in many organizations, authority structures do not evolve alongside that shift. Without explicit recalibration, responsibility is diffused, creating evaluation instability rather than empowerment.

### How Instability Surfaces

- Teams push more decisions upward when ownership is unclear.
- Managers spend more time correcting outputs than setting direction.
- Standards diverge across teams depending on comfort with AI.

What appears as productivity at the edge can become ambiguity at the center.

### Structural Examination

- Who owns the final decision when AI produces the first answer?
- Have performance standards been updated to distinguish speed from quality?
- Are escalation protocols calibrated for AI-assisted workflows?

### Leadership Imperative

Governance must evolve in parallel with capability. Clarify decision ownership before expanding AI's operational footprint.

## 3 Capacity Distortion

*Has efficiency translated into simplification—or escalation?*

AI makes work faster. But instead of doing less, teams are expected to do more. Output goes up, but the amount of work doesn't go down.

### How Instability Surfaces

- Output per team increases, but projects aren't moving faster.
- Leaders assume capacity has expanded—but no work has been taken off the table.
- Quality becomes less consistent as timelines are compressed.

The system produces more—without becoming lighter.

### Structural Examination

- Has AI reduced structural workload, or has it only expanded deliverables?
- Have timelines changed meaningfully?
- Is cost-to-serve declining in proportion to output growth?

### Leadership Imperative

Tie AI acceleration to deliberate workload decisions. Efficiency gains must produce structural relief, not permanent intensity.

## 4 Coordination Load

*Is acceleration increasing integration pressure?*

Usage of AI is rarely scaled uniformly. Adoption spreads across functions, creating new dependencies. If integration architecture does not evolve, coordination load expands. This is not a tooling issue. It is a systems issue.

### How Instability Surfaces

- Cross-functional synchronization effort increases.
- Outputs generated in one function require additional reconciliation in another.
- Enterprise initiatives stall due to integration friction.

Local acceleration creates system-wide slowdowns.

### Structural Examination

- Has coordination effort declined or intensified since AI deployment?
- Are integration standards consistently enforced?
- Is time-to-value being shortened or extended?

### Leadership Imperative

Strengthen orchestration before scaling. Innovation velocity must be matched by integration discipline.

# 5 Capability Integrity

*Is AI strengthening enterprise judgment—or compressing it?*

AI accelerates foundational tasks that historically built expertise. When those cycles are compressed without redesign, developmental pathways are narrowed. Recent research suggests that AI-assisted work can also increase cognitive fatigue when evaluation demands expand faster than workers' perceived control (*HBR*, 2026). The cognitive burden does not disappear; it simply is concentrated in fewer, more consequential decisions.

## How Instability Surfaces

- Greater reliance on AI-generated outputs without growth in independent analytical confidence
- Reduced coaching and formative feedback cycles
- Manager uncertainty when evaluating AI-assisted performance

Short-term productivity can coexist with long-term capability erosion.

## Structural Examination

- Are employees developing judgment capacity alongside AI use?
- Have learning pathways been redesigned for AI-assisted workflows?
- Do managers have clear standards for evaluating augmented work?

## Leadership Imperative

Treat human capability as infrastructure. Embed upskilling and evaluation discipline into AI-enabled processes before scaling further.

# **The Leadership Pivot:** Resetting the Human System Before Scaling

AI transformation rarely breaks because of insufficient capital or flawed tools. It falters when expectations accelerate while the underlying human system—roles, incentives, decision rights, communication norms—remains unchanged.

Early phases often show visible gains. But as AI expands, execution stability depends less on technological fluency and more on structural alignment. Durable transformation requires that:

- Roles evolve alongside capability
- Legacy work is eliminated, not layered
- Decision ownership is clarified
- Managers are equipped for new oversight demands
- Capacity trade-offs are made explicit

Scaling without resetting these foundations compounds fragility.

The most durable AI transformations are not those that deploy fastest. They are those that stabilize before expanding, reducing structural friction at the human layer before increasing acceleration at the technical layer. In short, **AI rewards organizations that stabilize before they scale.**

**CASE IN POINT:**

# When Structural Gaps Surface Early

A Fortune 500 company undertaking a large-scale digital transformation introduced new tools and new ways of working across regions and functions to modernize operations. Early indicators were encouraging. Deployment moved quickly. Adoption appeared steady.

What leadership lacked was visibility into how the transformation was functioning beneath those metrics:

- Did employees understand how their roles connected to the strategy?
- Were expectations consistent across regions?
- Was workload being reduced or simply redistributed?
- Were managers and teams equipped to reinforce the shift?

With the help of pattern intelligence, the organization examined how the transformation was experienced day to day. What surfaced were structural gaps that traditional surveys and dashboards had not revealed:

- Only **34%** of employees **clearly understood how their role connected** to the transformation strategy.
- **38%** reported **no positive sentiment** about transformation communications.
- **58%** reported **increased workload** without clarity on priorities or trade-offs.
- Only **slightly over half** of employees **reported feeling prepared** and engaged in the transition.

The constraint was not deployment. It was the organization's ability to absorb what had already been deployed. Role clarity, governance alignment, and workload discipline had not kept pace with change. Left unaddressed, these gaps would have compounded into coordination drag, rising oversight burden, and execution volatility.

**What made the difference was timing.** With the help of Seramount, leaders intervened before instability hardened, clarifying ownership, standardizing messaging, simplifying processes, and realigning expectations while adoption was still maturing. That window matters. Early detection is what allowed for continued acceleration without erosion. The longer instability goes unaddressed, the more expensive it becomes to reverse.

# **Conclusion:** Stabilize Before You Scale

AI transformation does not typically fail at the point of deployment. It destabilizes when structural adaptation lags behind acceleration. Execution strain surfaces early, often before financial indicators register decline. Output rises. Adoption appears strong. But beneath those signals, coordination load expands, managerial bandwidth tightens, and cost-to-serve begins to creep.

Most organizations are no longer deciding whether to adopt AI. They are already integrating it. The critical window is not at the beginning of transformation; it is during expansion. Structural misalignment compounds quietly and becomes exponentially more expensive to correct once scale hardens.

AI accelerates production. It does not automatically reset workflows, decision rights, coordination architecture, or capability pathways. Those shifts require deliberate intervention. The organizations that capture durable value from AI are not those that scale fastest. They are those that recognize execution strain early—and act before instability compounds.

**Without structural alignment, acceleration increases fragility. With it, AI's promised gains translate into competitive advantage.**

seramount

# The **Workforce** and **Talent Authority**

*Built on over 40 Years of Research*

## From Insight to Impact

*An integrated system of capabilities*

### Advisory and Research

Expert-led guidance that draws on original research and executive peer insight

### Benchmarking and Analytics

Market-defining global and sector benchmarks that turn data into direction

### Assessment and Strategy

Clear diagnostics that align leadership, speed execution, and activate lasting change

### Early-Talent Solutions

Technology and strategy that build resilient early-career pipelines

## Proven at Enterprise Scale

*Trusted by the world's leading organizations*

**600+**

**partners, 50%** of Fortune 100 and **25%** of Fortune 500

**5,000+**

**validated practices** informed by real-world application

**100+**

**annual convenings** designed to accelerate learning into action across leaders and their teams

# The AI Execution Risk Index

**Project Director**
Stephanie Larson, PhD

**Designer**
Kelsey Stoneham

seramount

Talent Solutions for Better Workplaces

202-747-1000 **seramount.com**