NEWROCKET

Executive White Paper

How to Go From Experimentation to Enterprise Impact



What it Takes to Scale and Transform with AI

Frank Palermo Chief Operating Officer

Copyright © 2026 NewRocket LLC.

Table of Contents

01	Executive Summary	03
02	From Experimentation to Enterprise Impact	04
03	We Don’t Have an AI Experimentation Problem	05
04	We Have an AI Scaling Problem	06
05	AI Value Compounds as You Move from Individual to Enterprise	08
06	The Model Is Rarely the Hardest Part	09
07	Stop Adding AI to Old Processes: Redesign the Work	10
08	Agents Change the Operating Model	12
09	Governance Must Move at the Speed of AI	13
10	Enterprise AI Needs an Operating System for Scale	14
11	The Constraint Is Shifting from Technology to Leadership	16
12	Is Your Enterprise Ready to Operate Differently?	18
13	Executive Action Agenda	20
14	Sources and Notes	21

01

Executive Summary



Frank Palermo
Chief Operating Officer, NewRocket

October 2026

> *The companies that capture the greatest value from AI will not necessarily be those with the best models or the most agents. They will be the organizations most willing and able to operate differently.*

Enterprise AI has reached an inflection point. The technology is no longer waiting for organizations to prove that it can generate text, summarize documents, write code, analyze data, or assist employees. Those capabilities are increasingly accessible and increasingly embedded in the tools people already use. The harder question is whether enterprises can convert rapidly improving intelligence into durable operating advantage.

The evidence points to a widening gap between adoption and value. Stanford's 2025 AI Index reported that 78% of organizations used AI in 2024, up from 55% a year earlier. By 2025, McKinsey reported AI use in at least one business function at 88%, yet only 7% of survey respondents said AI had been fully scaled across their organizations. BCG's 2025 research was even more pointed: only 5% of companies in its global study were achieving AI value at scale, while 60% reported little material value despite significant investment.[1][2][3]

This gap exists because enterprise AI is changing categories. The first wave was primarily about assistance: helping an individual produce an answer, draft content, analyze information, or write code. The next wave is about agency: AI that can reason over enterprise context, select tools, make bounded decisions, take action in systems, coordinate work, and escalate exceptions. That transition changes the management problem. The question is no longer only whether a model is accurate enough. It is whether the enterprise has designed the context, workflow, authority, controls, accountability, and measurement required for AI to participate in real work.

The central argument of this paper is simple: the companies that capture the greatest value from AI will not necessarily be those with the best models or the most agents. They will be the organizations most willing and able to operate differently. That requires moving beyond isolated productivity gains toward redesigned end-to-end work, human-agent operating models, embedded governance, and measurable business outcomes.

The opportunity is substantial because AI can compress the cost of cognition inside the enterprise. Activities that historically required people to search, interpret, reconcile, draft, decide, route, and follow up can increasingly be augmented or executed by software. But

lower-cost cognition does not automatically create lower-cost operations. If organizations preserve the same queues, handoffs, approval layers, organizational boundaries, and metrics, much of the potential value is absorbed by the existing operating model. Transformation begins when leaders redesign the system around what is newly possible.

> The next phase of enterprise AI will be won by organizations that redesign how work gets done—not by organizations that simply add more AI to how work gets done today.

02

From Experimentation to Enterprise Impact

Over the last several years, AI has moved through three distinct enterprise conversations: experiment, scale, and transform. The first was about proving the technology. Could a large language model answer questions accurately enough? Could it summarize a contract, assist a developer, triage a service request, or generate useful content? The answer, increasingly, is yes.

The second conversation is about scale. Can those capabilities operate reliably across thousands of employees, millions of transactions, multiple systems, sensitive data, and complex regulatory environments? Can they perform consistently when the data is incomplete, permissions differ by user, policies change, exceptions occur, and downstream actions have financial or customer consequences? This is where many organizations are now struggling.

The third conversation—transformation—is more consequential still. It asks what happens when AI is not simply another tool used inside the current enterprise, but becomes an active participant in how work is performed. An AI assistant that helps an employee draft a response may create productivity. An agent that understands context, accesses enterprise systems, determines the next action, executes it, and escalates only exceptions changes the economics and design of the process itself.

These stages require different leadership disciplines. Experimentation rewards speed,

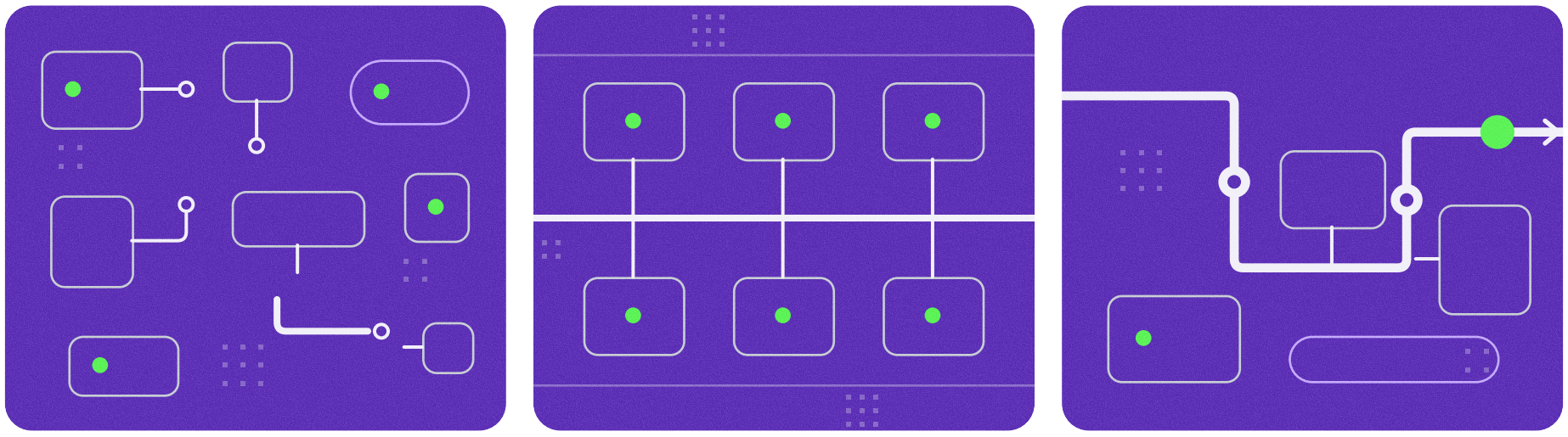


Figure adapted from the final T200 SOAR presentation: From Experimentation to Enterprise Impact.

curiosity, and technical learning. Scale requires engineering discipline, controls, integration, support, and repeatability. Transformation requires operating-model choices: what work should disappear, what work should be redesigned, what authority should move to software, how roles should change, and how value should be captured. Organizations often struggle because they attempt to solve a transformation problem with an experimentation operating model.

Enterprise impact therefore should not be defined as "more AI." It should be defined as measurable change in business performance: lower cost-to-serve, faster cycle times, higher quality, better customer or employee experience, reduced risk, increased revenue capacity, or new products and services that were previously uneconomic. The progression from experiment to scale to transform is ultimately a progression from proving capability to changing enterprise economics.

03

We Don't Have an AI Experimentation Problem

Most large organizations do not lack AI activity. They have copilots, pilots, hackathons, internal assistants, agents, proofs of concept, and long inventories of proposed use cases. AI can be found in contract review, sales research, demand forecasting, claims intake, service triage, software engineering, order management, IT operations, HR service delivery, and dozens of other domains.

This activity is understandable. The potential is enormous. McKinsey estimates that current

generative AI and other technologies could theoretically automate activities that consume 60% to 70% of employees' working time. Its research also estimates that roughly 75% of generative AI's potential value is concentrated in customer operations, marketing and sales, software engineering, and R&D.[4]

But experimentation can create a dangerous illusion of progress. A company can have hundreds of AI use cases and still leave its core operating model largely unchanged. During a recent executive AI workshop, a CEO posed a deceptively simple question to the leadership team: "We are spending an incredible amount on AI. What has actually changed in the business?" The responses described adoption inside individual functions and the number of use cases underway. What was missing was evidence that an end-to-end business process had been redesigned, accountability had changed, or a measurable enterprise outcome had materially improved.

This is a common pattern because use-case portfolios naturally fragment ownership. Each function finds opportunities inside its own boundaries. Finance automates finance tasks. HR improves HR service. Sales adds sales assistants. IT introduces coding and operations agents. Local value is real, but the organization can optimize dozens of tasks without changing the cross-functional flow of work that customers and employees actually experience.

Leaders therefore need two scoreboards. The first measures adoption and learning: active users, use cases, pilot velocity, model quality, and employee engagement. The second measures transformation: end-to-end cycle time, cost per outcome, straight-through processing, exception rate, quality, revenue conversion, customer effort, employee capacity released, and realized financial impact. The first tells you whether AI is being used. The second tells you whether the business is changing.

The distinction is crucial. The number of pilots is an input metric. AI transformation begins when the enterprise can point to a process and say: this now operates differently, produces a better outcome, and has a new economic profile because AI is part of the operating model.

04

We Have an AI Scaling Problem

The problem begins after the demo. A proof of concept can be built around curated data, limited users, narrow permissions, manual supervision, and a carefully selected workflow. Production systems do not have that luxury. They must operate across messy data, inconsistent processes, changing policies, multiple platforms, security boundaries, organizational silos, and real customers.

This is why the funnel from ideas to scaled capability narrows so dramatically. The presentation that inspired this paper illustrated the pattern as 200 ideas becoming 50 pilots, 10



Figure adapted from the final T200 SOAR presentation: We Have an AI Scaling Problem.

production deployments, and perhaps two capabilities operating at true enterprise scale. Those figures are illustrative, but the broader market data points in the same direction. McKinsey's 2025 survey found widespread AI use but only 7% of respondents reporting full organizational scale. BCG found just 5% of companies achieving AI value at scale.[2][3]

Scaling fails for predictable reasons. Data that looked sufficient in a pilot is incomplete or inaccessible in production. The workflow depends on exceptions that were not modeled. The agent needs permissions that security teams are unwilling to grant. The process crosses systems owned by different groups. No one owns the end-to-end outcome. Risk teams cannot see what the agent did. Employees do not trust the recommendation. The economics disappear once integration, supervision, support, and exception handling are included.

There is also an architectural transition. A single model call is relatively simple. A production agent may need retrieval, memory, identity, tool use, policy checks, deterministic business logic, event triggers, human approval, observability, evaluation, and fallback behavior. The reliability of the system is determined by the interaction of all these components—not by the model benchmark alone.

For leaders, this changes portfolio management. Rather than funding the largest possible list of pilots, organizations should deliberately select a smaller number of value streams where they are prepared to solve the surrounding enterprise problems. The objective is to build reusable capabilities—context, integration, governance, evaluation, orchestration—not just isolated use cases.

The implication is important: organizations should stop treating the scaling gap as a temporary technical inconvenience. It is the primary management problem of enterprise AI. Scaling

requires a repeatable way to connect AI to enterprise context, integrate it into workflows, govern what it can do, manage adoption, and measure whether the economics are real.

05

AI Value Compounds as You Move from Individual to Enterprise

AI value can be viewed as a progression. At the first level, individual productivity, employees use AI to write, summarize, research, analyze, or code more efficiently. The enablers are relatively straightforward: tools, training, context, and adoption. These gains matter because they improve the economics of knowledge work and create familiarity with AI.

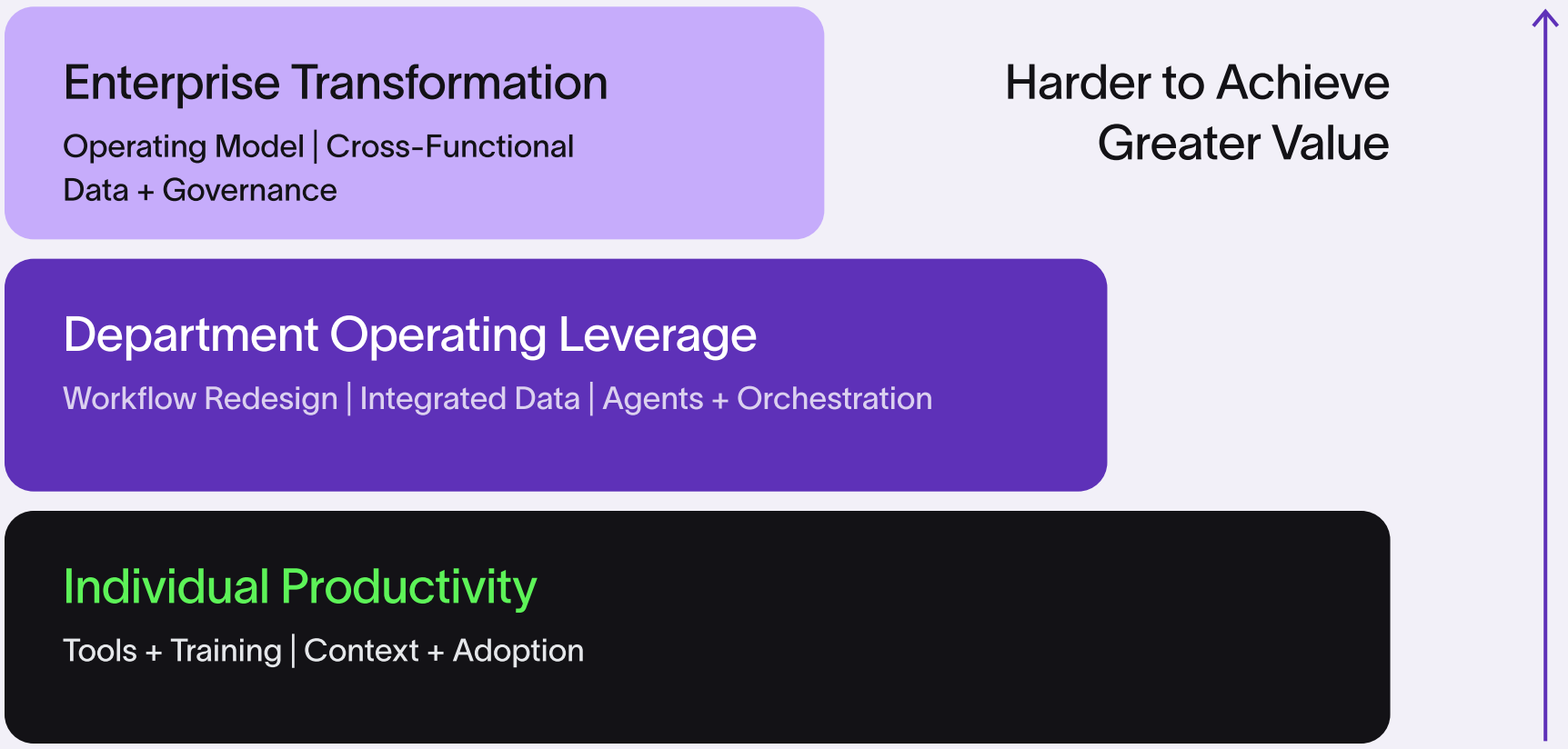


Figure adapted from the final T200 SOAR presentation: AI Value Compounds as You Move from Individual to Enterprise.

The next level is departmental operating leverage. Here, AI becomes part of a workflow rather than a personal tool. The unit of change shifts from the task to the process. Achieving value requires integrated data, workflow redesign, orchestration, process ownership, and a clearer understanding of where humans and agents each contribute.

At the enterprise level, the challenge becomes cross-functional. Customer onboarding, lead-to-cash, order-to-activate, incident-to-resolution, hire-to-retire, and procure-to-pay rarely belong to one department. Data, processes, controls, ownership, and incentives must cross organizational boundaries. That is harder—but the potential value is greater because the organization is changing an entire value stream rather than accelerating isolated tasks.

The economics also change. Individual productivity often creates "soft" capacity: minutes saved across many people. Unless leaders redesign roles, workload, staffing, or service levels, that capacity may not appear as a financial outcome. Workflow redesign makes the value more capturable. It can remove queues, eliminate rework, reduce the number of touches, increase throughput, and allow people to focus on exceptions or higher-value work.

This suggests a useful hierarchy of AI metrics. At the individual level, measure time saved and quality improvement. At the workflow level, measure touches, cycle time, exception rates, straight-through processing, and service levels. At the enterprise level, measure unit economics, revenue, risk, customer experience, workforce capacity, and capital efficiency. The metric should rise with the level of transformation.

The value therefore compounds as AI moves from person, to department, to enterprise. But so does the transformation required to capture it.
The mistake is assuming that enough individual productivity improvements will automatically add up to enterprise transformation. They usually do not.

Enterprise value must be designed.

06

The Model Is Rarely the Hardest Part

Model capability receives a disproportionate amount of attention because it is visible and advancing quickly. In production, however, the model is only one component of a much larger system. The barriers surrounding it are often more difficult: fragmented data, disconnected workflows, integration, identity and access, process ownership, risk, adoption, governance, and accountability.

BCG's long-standing 10-20-70 framework is useful here. It argues that roughly 10% of AI transformation effort should focus on algorithms, 20% on technology and data, and 70% on people and processes. BCG's more recent work on scaling AI reiterates the same point: companies should devote the majority of their effort to people and processes if they want AI-driven changes to stick.[5]

Context is particularly important. An enterprise agent needs more than a prompt. It needs access to the right policies, customer history, product data, workflow state, prior decisions, organizational roles, entitlements, and system events at the moment it acts. This "context architecture" becomes a critical enterprise capability: determining what the AI needs to know, what it is permitted to know, how that context is assembled, how fresh it is, and how its quality and provenance are maintained.

Context architecture is broader than retrieval-augmented generation. Retrieval answers the question, "What information should the model see?" Enterprise context also answers, "Who is asking? What role do they have? What process are we in? What happened before? What systems are authoritative? What policies apply? What actions are allowed? What evidence must be retained?" Without these dimensions, an agent can be knowledgeable and still be operationally unsafe or ineffective.

Integration is equally strategic. Intelligence creates value only when it can affect work. That means connecting AI to systems of record, workflow platforms, APIs, event streams, data platforms, communication channels, and tools. Organizations that treat integration as a project-by-project afterthought will recreate the same fragmentation that limited earlier automation waves.

Finally, people and accountability matter because AI changes decision rights. A process owner must decide what can be automated, which exceptions require escalation, what level of error is tolerable, how performance is monitored, and who is accountable when the system produces an undesirable outcome. These are business design decisions, not model configuration settings.

As models improve and model choice becomes more fluid, durable competitive advantage will increasingly come from the system around the model: proprietary context, redesigned workflows, integration, governance, and the ability to turn intelligence into action.

07

Stop Adding AI to Old Processes: Redesign the Work

One of the most important shifts leaders can make is changing the question. Instead of asking, "Where can we add AI to this process?" ask: "If AI were part of the workforce, how would we design this process?" The first question encourages organizations to bolt AI onto existing work.

A human performs a step, an AI assistant helps with the next step, the work enters a queue, another employee reviews it, and the process continues. Some tasks get faster, but the architecture of the process remains intact—including its handoffs, controls, delays, and organizational boundaries.

The second question starts from first principles. What outcome is the process intended to produce? Which decisions require human judgment? Which are rules-based or sufficiently bounded for an agent? Which handoffs exist because systems historically could not reason across domains? Which reviews are necessary risk controls, and which are simply artifacts of uncertainty? Where should exceptions trigger human intervention rather than forcing humans to process every transaction?

This is where AI begins to challenge conventional process design. Traditional workflows were often optimized around the limitations of people and software: people could interpret ambiguity but were expensive and capacity constrained; software could execute quickly but required deterministic rules. Agentic AI introduces a third design option—software that can interpret context and reason within boundaries. That allows organizations to reconsider where work waits, where it moves, and where decisions happen.

A redesigned process may look radically simpler. Instead of human-to-system-to-queue-to-human-to-queue-to-human, the flow can become trigger-to-agent-to-human-checkpoint-to-complete.

The human checkpoint is not removed; it is repositioned where judgment, accountability, or risk actually requires it. The goal is not to eliminate humans. It is to eliminate unnecessary human handling.

McKinsey's research reinforces this point: among the organizational practices it studied, workflow redesign had the largest effect on an organization's ability to see EBIT impact from generative AI. Its guidance for corporate functions similarly argues that organizations should "rewire ways of working" rather than merely automate tasks.[6]

A practical redesign method begins with the outcome, maps the decisions required to produce it, classifies those decisions by risk and ambiguity, defines the context required for each decision, and then assigns the work to humans, deterministic automation, or agents. Only after that should the organization design the workflow and technology.

The target state is not a fully autonomous enterprise. It is a deliberately redesigned enterprise in which people and AI each perform the work for which they are best suited. The design principle is not maximum automation. It is maximum business value within acceptable risk.

If AI were part of the workforce,
how would we design this process?



08

Agents Change the Operating Model

The operating-model question becomes unavoidable as AI evolves from generating and recommending to taking action. Once an agent can access systems, make decisions, invoke tools, coordinate with other agents, and execute work, leaders must define authority as deliberately as they define technology.

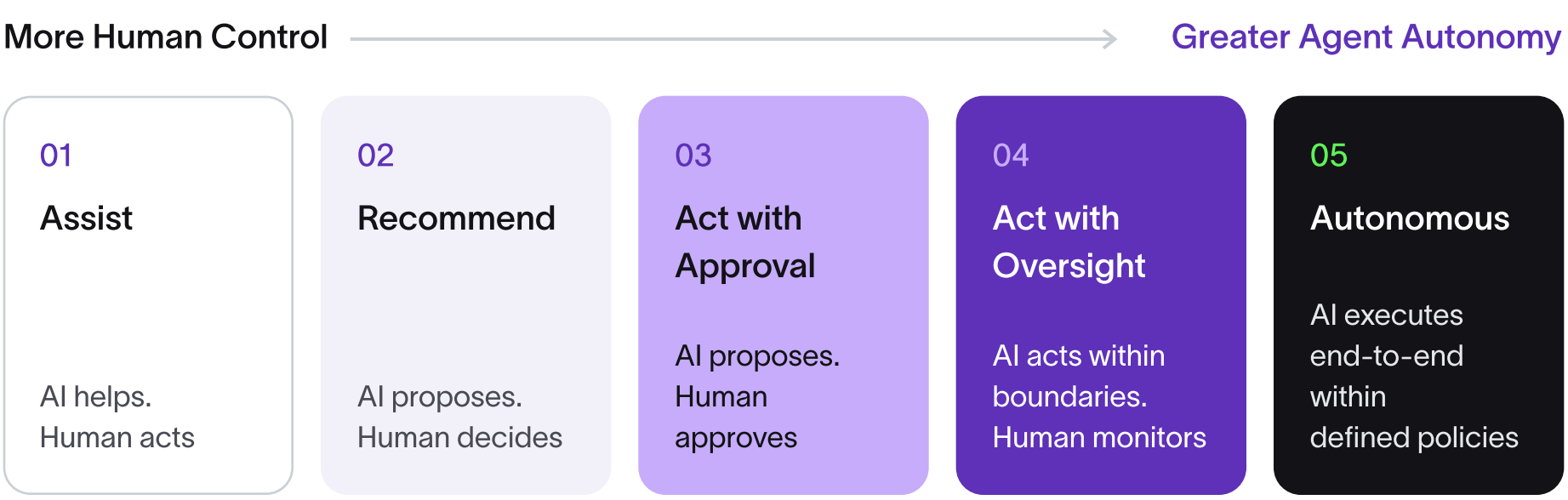


Figure adapted from the final T200 SOAR presentation: Agents Change the Operating Model.

A useful autonomy continuum is: Assist; Recommend; Act with Approval; Act with Oversight; Autonomous. Different activities should occupy different points on that spectrum. An agent may autonomously classify a low-risk service request, recommend a response for a sensitive employee issue, act with approval on a financial exception, or execute an end-to-end workflow inside clearly defined policy boundaries.

The correct level of autonomy depends on more than technical confidence. It depends on the consequence of error, reversibility of the action, quality of context, regulatory requirements, customer sensitivity, model and workflow performance, and the organization's ability to observe what happened. A reversible low-value action can tolerate more autonomy than an irreversible high-value decision.

This creates four management questions. Who decides? Who acts? When do humans intervene? Who is accountable? A fifth question—how do we observe it?—becomes essential once agents operate at machine speed and across systems.

The governing principle is straightforward: authority can move to the agent; accountability

cannot. The organization still needs a named human owner for the process, policy, risk thresholds, performance, and outcome. Autonomy without accountability is not an operating model; it is unmanaged delegation.

This also changes jobs. In an agent-enabled process, human work shifts toward defining goals, handling exceptions, exercising judgment, supervising performance, improving policies, resolving ambiguous cases, and redesigning the system. Some roles will become more leveraged; others will narrow or combine. Managers will increasingly manage a blended capacity model of people, automation, and agents rather than people alone.

Microsoft's 2025 Work Trend Index points toward this emerging model. In a survey and analysis spanning 31,000 workers in 31 countries, 81% of leaders said they expected agents to be moderately or extensively integrated into their AI strategy within 12 to 18 months, while 82% said the year was pivotal for rethinking strategy and operations.[7]

The organizational challenge is to make autonomy explicit rather than accidental. Every agentic process should have a documented autonomy level, allowed actions, required approvals, escalation triggers, owner, performance measures, and conditions under which autonomy can be increased—or reduced.

✅ Authority can move to the agent.

❌ Accountability cannot.

09

Governance Must Move at the Speed of AI

Traditional governance often assumes that technology is relatively static: approve a system, establish controls, audit periodically, and manage exceptions. Agentic AI changes that assumption. Agents can make sequential decisions, invoke tools, interact with other systems, and produce actions whose risk depends on context.

Governance therefore has to move from a periodic review activity to an embedded operating capability.

A practical governance model begins by knowing what exists: agents, models, third-party AI, and shadow AI. Discovery and inventory are foundational because organizations cannot govern assets they cannot see. Inventory also needs ownership. Every material AI asset should have a named business and technical owner, a defined purpose, a risk classification, and an understood lifecycle.

The next requirement is observability. Traditional application monitoring tells an organization whether a system is up. Agentic observability must help explain what the system decided, which context it used, which tools it invoked, which policies were applied, what action it took, and what happened next. Traces, evaluations, alerts, cost telemetry, and outcome metrics become part of the control environment.

Control must operate at runtime. Identity and access should be scoped to least privilege. Policies and guardrails should define allowed and prohibited actions. High-risk actions may require approval. Runtime enforcement should be able to detect, block, or stop behavior that exceeds policy. Reversibility matters because the ability to trace and undo an action can materially change the acceptable autonomy level.

Governance must also prove conformance and value. Regulatory mapping, policy evidence, audit trails, model and agent evaluations, and lifecycle records support compliance.

At the same time, organizations should govern economics: usage, token or compute spend, adoption, value realization, and cost caps. An AI asset that is safe but creates no value is still a governance problem.

NIST's AI Risk Management Framework supports this lifecycle view. Its four core functions—Govern, Map, Measure, and Manage—are intended to make AI risk management continuous and embedded across the AI lifecycle. NIST explicitly describes governance as a cross-cutting function rather than a one-time gate.[8]

The goal is not to create a governance bureaucracy that slows AI to the pace of a committee. The goal is to make safe behavior part of the architecture: identity, policy, controls, observability, evidence, escalation, and reversibility built into the way AI operates. Effective governance should increase the organization's confidence to scale, not simply constrain it.

10

Enterprise AI Needs an Operating System for Scale

Scaling AI requires more than AI. It requires a coherent operating system that brings together five capabilities: trusted data, redesigned work, agents, governance, and outcomes. These are not separate workstreams; they are mutually dependent elements of a production enterprise AI system.

Trusted Data provides context, quality, access, integration, provenance, and the semantic grounding required for AI to act on enterprise

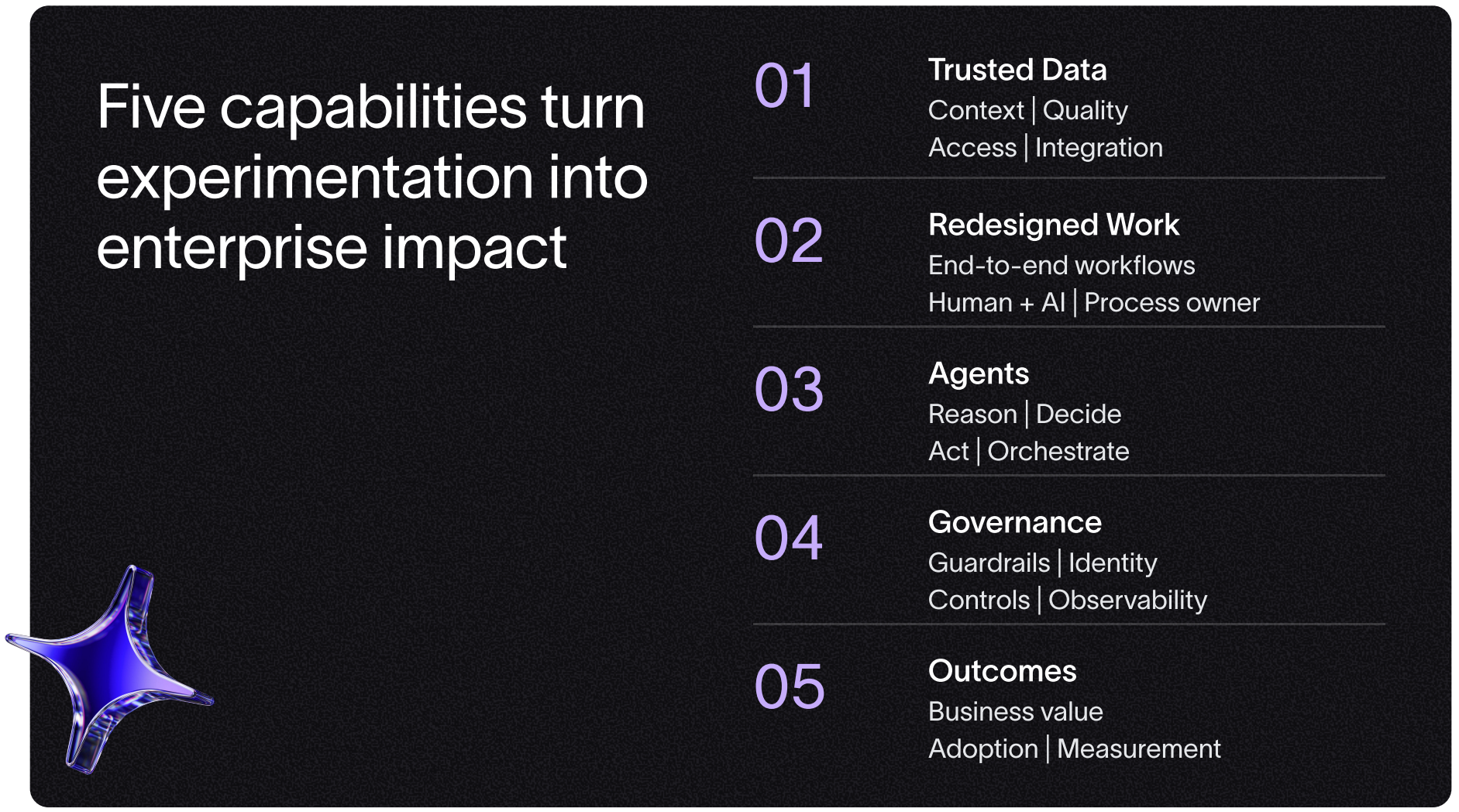


Figure adapted from the final T200 SOAR presentation: Enterprise AI Needs an Operating System for Scale.

reality. The issue is not simply whether data exists. It is whether the right data can be assembled at the moment of decision, with the correct identity and permissions, from authoritative sources, in a form the agent can use.

Redesigned Work defines the end-to-end workflow, the role of humans and AI, decision points, exceptions, and process ownership. Without redesign, AI tends to optimize tasks inside the existing process. With redesign, the organization can remove steps, collapse handoffs, reposition controls, and create straight-through execution where risk allows.

Agents provide reasoning, decisioning, action, and orchestration. They are the execution layer that can interpret context and move work forward. But an agent is valuable only when it is connected to the right context and tools, placed inside a redesigned workflow, and governed according to the consequences of its actions.

Governance provides guardrails, identity, controls, observability, evaluation, lifecycle management, and evidence. It establishes the boundaries within which agents can operate and the mechanisms through which humans can understand, intervene, and remain accountable.

Outcomes close the loop. Business value, adoption, service levels, quality, risk, and financial measures determine whether the AI system is producing the intended result. Outcome telemetry should feed continuous improvement: where performance is strong, autonomy can expand; where exceptions or errors rise, the workflow, context, policy, or model should be adjusted.

The sequence is important conceptually, even if implementation is iterative. Trusted data gives AI

the context it needs. Redesigned work determines where intelligence should be applied. Agents convert intelligence into decisions and action. Governance establishes boundaries and evidence. Outcomes create the feedback loop that tells the organization whether the system is actually producing value.

Many AI programs overinvest in the center of this model—the agent—and underinvest in everything around it. That is why an impressive demonstration can coexist with weak enterprise economics. The agent is not the operating system. It is one component of it.

The most mature organizations will turn these five capabilities into shared enterprise services rather than rebuilding them for every use case. Reusable context services, identity patterns, integration connectors, evaluation harnesses, policy controls, observability, human-in-the-loop patterns, and value measurement can dramatically reduce the marginal cost and risk of deploying the next AI-enabled workflow.

11

The Constraint Is Shifting from Technology to Leadership

AI capability is advancing faster than most enterprises can change their operating models. Models can already generate, summarize, analyze, recommend, use tools, take action, and increasingly coordinate multi-step work. The organizational response is slower. Enterprises must enable tools and access, redesign workflows, redefine human-AI collaboration, and ultimately transform roles, accountability, and operating structures.

That creates an enterprise readiness gap: the difference between what AI can technically do and what the organization is prepared to let it do responsibly and productively. The gap can widen even as technology improves. A more capable model does not automatically make the enterprise more capable of deploying it.

Closing that gap is a leadership responsibility. It requires choices that cannot be delegated to the AI team: Which value streams matter most? Which processes should be redesigned? What decisions can agents make? Where must human judgment remain? Who owns the end-to-end outcome? What risks are acceptable? How will productivity gains be captured rather than merely absorbed? How will roles, spans, skills, and incentives change?

Leadership also has to resolve structural friction. End-to-end AI transformation often crosses the boundaries of the organization chart. A lead-to-cash redesign may involve marketing, sales, legal, finance, operations, and technology. If every function optimizes its own step but no executive owns the entire value stream, AI will reproduce the fragmentation already present in the enterprise.

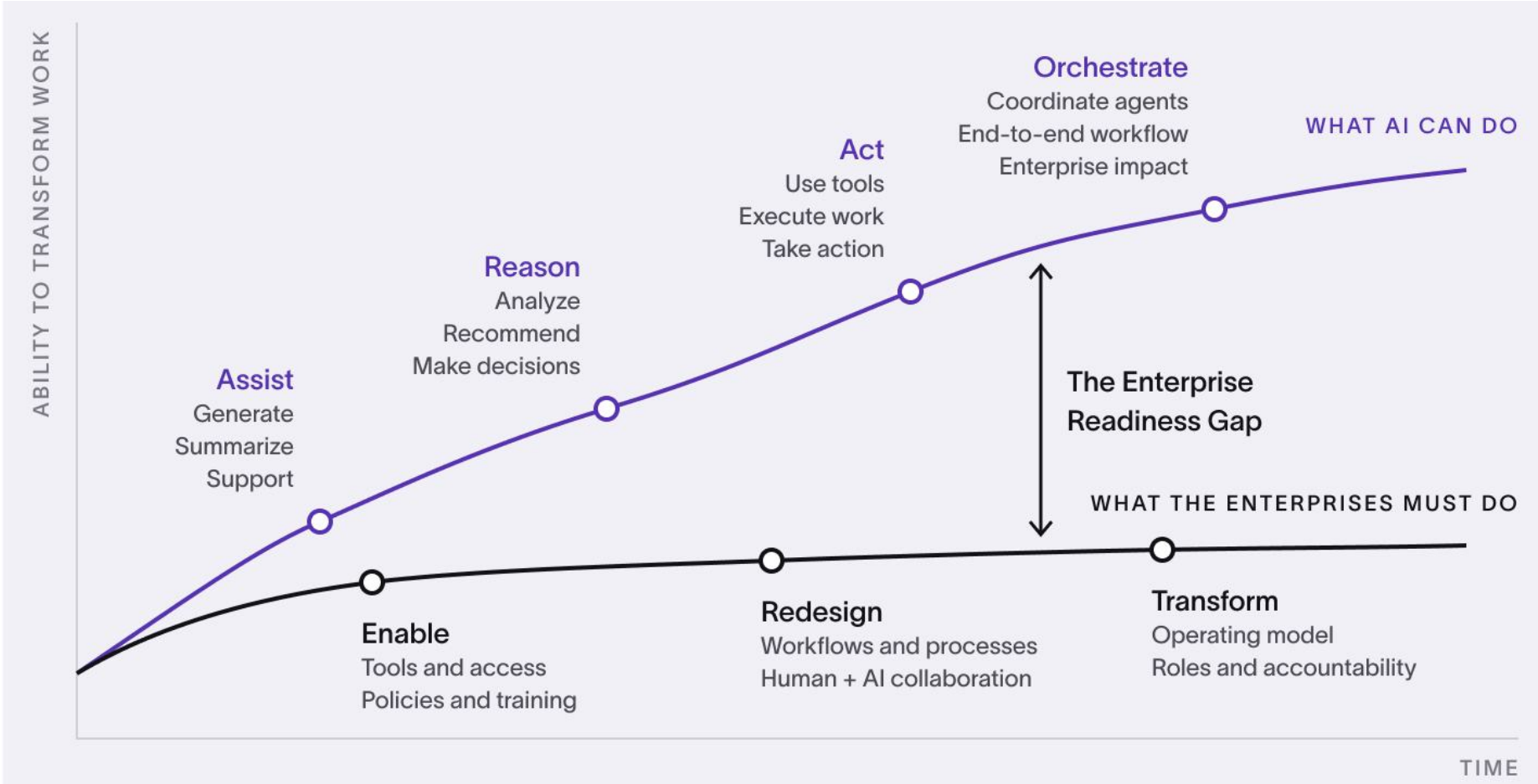


Figure adapted from the final T200 SOAR presentation: The Constraint Is Shifting from Technology to Leadership.

The personal leadership traits matter as well. Curiosity is needed to understand what is newly possible. Courage is required to challenge processes and organizational boundaries that have existed for years. Decisiveness is necessary because waiting for the technology to stabilize is not a strategy. Systems thinking is essential because local optimization can create enterprise complexity. Adaptability matters because both the technology and the operating model will evolve continuously.

Leaders also need discernment. The presence of a capable agent does not mean every decision should be delegated. The best AI leaders will distinguish between work that should be assisted, work that should be automated, work that should be redesigned, and work that should remain primarily human. They will treat autonomy as a management variable rather than an ideology.

BCG's research captures the underlying issue well: leading organizations devote the majority of AI transformation effort to people and processes. The technology is moving quickly. Leadership capacity to redesign the enterprise is increasingly the scarce resource.[5]

12

Is Your Enterprise Ready to Operate Differently?



The AI-native enterprise will not simply be today's organization with better software. It will operate differently. Human-led work will increasingly become human-agent work. Functional silos will give way, where appropriate, to cross-functional orchestration. Manual handoffs will be replaced by context-rich workflows. AI will move from assisting to acting inside defined boundaries. Humans will shift from executing every step to designing, governing, supervising, handling exceptions, and remaining accountable for outcomes.

The transition will not happen everywhere at once, nor should it. Different processes have different economics, risk profiles, regulatory requirements, and tolerance for autonomy. The objective is not to race toward maximum autonomy. It is to build the organizational capability to place each process at the appropriate level of autonomy and then increase that level as evidence, trust, and controls mature.

Readiness can be assessed across six dimensions. Strategy asks whether the organization has selected business outcomes and value streams rather than simply collected use cases. Context asks whether AI can access trusted, permissioned enterprise information. Work design asks whether processes have been redesigned from first principles. Operating model asks whether human and agent roles, decision rights, ownership, and escalation are explicit. Governance asks whether the organization can see, control, evaluate, and reverse agent behavior. Value asks whether adoption and technical performance connect to measurable business outcomes.

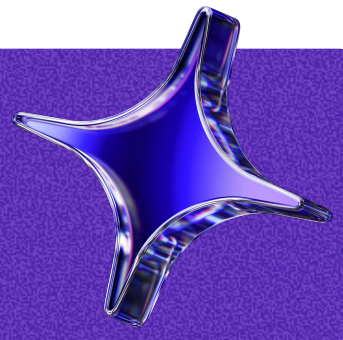
Executives can use these dimensions to distinguish enthusiasm from readiness. A company may be advanced in model experimentation but immature in context. It may have excellent governance policies but weak workflow redesign. It may deploy agents successfully but lack a mechanism for capturing productivity gains. The weakest capability can become the constraint on enterprise scale.

The practical goal is to create a transformation portfolio, not just an AI portfolio. Select a small number of cross-functional value streams, establish an executive owner, baseline the economics, redesign the work, define the target autonomy model, build the required context and controls, deploy incrementally, and measure the

result. Then use what is learned to create reusable enterprise capabilities for the next value stream.

The organizations that do this well will be positioned to cross the gap from experimentation to enterprise impact. Those that do not may still deploy impressive AI—but they risk accumulating more tools, more pilots, and more agents without fundamentally changing the business.

The next competitive advantage in AI will not come from access to intelligence alone. Intelligence is becoming abundant. Advantage will come from the ability to reorganize the enterprise around it.



The question is no longer whether AI is ready for the enterprise. Increasingly, the question is whether the enterprise is ready to operate differently.

13

Executive Action Agenda

For leadership teams moving from AI experimentation toward scaled impact, the immediate task is to turn the ideas in this paper into a managed transformation portfolio. Five actions provide a practical starting point.

01

Concentrate on value streams, not use-case volume.

Select a small number of end-to-end processes where AI can materially change cost, speed, quality, growth, risk, or customer experience. Establish an executive owner and baseline the economics before redesign begins.

02

Redesign before automating.

Map the target operating model from first principles, including decisions, handoffs, exceptions, controls, and human-agent responsibilities. Remove unnecessary steps rather than accelerating them.

03

Build the context and control plane.

Treat enterprise data, semantics, identity, permissions, observability, policy, integration, evaluation, and reversibility as shared AI infrastructure—not one-off requirements for each pilot.

04

Define autonomy and accountability explicitly.

Choose the appropriate autonomy level for each decision based on consequence, reversibility, context quality, and control maturity. Maintain named human ownership for outcomes.

05

Measure business outcomes continuously.

Track adoption and technical performance, but govern investment around economic and operating results. Use outcome telemetry to decide where to expand autonomy, redesign the workflow, or stop the investment.

14

Sources and Notes

[1] Stanford Institute for Human-Centered Artificial Intelligence, The 2025 AI Index Report. Reports 78% of organizations using AI in 2024, up from 55% in 2023.

[2] McKinsey & Company, “AI at work but not at scale,” December 10, 2025, based on the 2025 McKinsey Global Survey on AI. Reports 88% using AI in at least one function and 7% fully scaled across organizations.

[3] Boston Consulting Group, The Widening AI Value Gap / 2025 AI value research. Study of more than 1,250 firms worldwide; 5% achieving AI value at scale, 35% scaling and generating some value, and 60% reporting minimal material value.

[4] McKinsey Global Institute, “The economic potential of generative AI: The next productivity frontier,” 2023. Estimates that current generative AI and other technologies could theoretically automate activities absorbing 60–70% of employees’ time; approximately 75% of generative AI value is concentrated in four functions.

[5] Boston Consulting Group, “The Leader’s Guide to Transforming with AI” and subsequent scaling research. BCG’s 10-20-70 principle allocates 10% of transformation effort to algorithms, 20% to technology and data, and 70% to people and processes.

[6] McKinsey & Company research on enterprise generative AI and corporate functions. McKinsey identifies workflow redesign as a major organizational practice associated with EBIT impact and recommends rewiring ways of working rather than simply automating tasks.

[7] Microsoft, 2025 Work Trend Index Annual Report, “The year the Frontier Firm is born,” April 23, 2025. Based on 31,000 workers across 31 countries plus Microsoft 365 and LinkedIn signals; 82% of leaders called 2025 pivotal for rethinking strategy and operations, and 81% expected agents to be moderately or extensively integrated into AI strategy within 12–18 months.

[8] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0) and AI RMF Playbook. Core functions: Govern, Map, Measure, and Manage.

[9] NewRocket, “From Experimentation to Enterprise Impact: What It Takes to Scale and Transform with AI,” T200 SOAR, September 15, 2026. Final presentation framework and operating-model concepts adapted throughout this draft.