The GenAI Divide

STATE OF AI IN BUSINESS 2025

MIT NANDA

Aditya Challapally
Chris Pease
Ramesh Raskar
Pradyumna Chari
July 2025

NOTES

Preliminary Findings from AI Implementation Research from Project NANDA
Reviewers: Pradyumna Chari, Project NANDA
Research Period: January – June 2025
Methodology: This report is based on a multi-method research design that includes a systematic review of over 300 publicly disclosed AI initiatives, structured interviews with representatives from 52 organizations, and survey responses from 153 senior leaders collected across four major industry conferences.
Disclaimer: The views expressed in this report are solely those of the authors and reviewers and do not reflect the positions of any affiliated employers.
Confidentiality Note: All company-specific data and quotes have been anonymized to maintain compliance with corporate disclosure policies and confidentiality agreements, ensure neutrality, and prevent any perception of commercial advancement or opinion.

1 CONTENTS

2 EXECUTIVE SUMMARY

Despite $30–40 billion in enterprise investment into GenAI, this report uncovers a surprising result in that 95% of organizations are getting zero return. The outcomes are so starkly divided across both buyers (enterprises, mid-market, SMBs) and builders (startups, vendors, consultancies) that we call it the GenAI Divide. Just 5% of integrated AI pilots are extracting millions in value, while the vast majority remain stuck with no measurable P&L impact. This divide does not seem to be driven by model quality or regulation, but seems to be determined by approach.

Tools like ChatGPT and Copilot are widely adopted. Over 80 percent of organizations have explored or piloted them, and nearly 40 percent report deployment. But these tools primarily enhance individual productivity, not P&L performance. Meanwhile, enterprise-grade systems, custom or vendor-sold, are being quietly rejected. Sixty percent of organizations evaluated such tools, but only 20 percent reached pilot stage and just 5 percent reached production. Most fail due to brittle workflows, lack of contextual learning, and misalignment with day-to-day operations.

From our interviews, surveys, and analysis of 300 public implementations, four patterns emerged that define the GenAI Divide:

- **Limited disruption**: Only 2 of 8 major sectors show meaningful structural change
- **Enterprise paradox**: Big firms lead in pilot volume but lag in scale-up
- **Investment bias**: Budgets favor visible, top-line functions over high-ROI back office
- **Implementation advantage**: External partnerships see twice the success rate of internal builds

The core barrier to scaling is not infrastructure, regulation, or talent. It is learning. Most GenAI systems do not retain feedback, adapt to context, or improve over time.

A small group of vendors and buyers are achieving faster progress by addressing these limitations directly. Buyers who succeed demand process-specific customization and evaluate tools based on business outcomes rather than software benchmarks. They expect

systems that integrate with existing processes and improve over time. Vendors meeting these expectations are securing multi-million-dollar deployments within months.

While most implementations don't drive headcount reduction, organizations that have crossed the GenAI Divide are beginning to see selective workforce impacts in customer support, software engineering, and administrative functions. In addition, the highest-performing organizations report measurable savings from reduced BPO spending and external agency use, particularly in back-office operations. Others cite improved customer retention and sales conversion through automated outreach and intelligent follow-up systems. These early results suggest that learning-capable systems, when targeted at specific processes, can deliver real value, even without major organizational restructuring.

3 The Wrong Side of the GenAI Divide: High Adoption, Low Transformation

Takeaway: Most organizations fall on the wrong side of the GenAI Divide, adoption is high, but disruption is low. Seven of nine sectors show little structural change. Enterprises are piloting GenAI tools, but very few reach deployment. Generic tools like ChatGPT are widely used, but custom solutions stall due to integration complexity and lack of fit with existing workflows.

The GenAI Divide is most visible when examining industry-level transformation patterns. Despite high-profile investment and widespread pilot activity, only a small fraction of organizations have moved beyond experimentation to achieve meaningful business transformation.

3.1 The Disruption Reality Behind the Divide

Takeaway: The GenAI Divide manifests clearly at the industry level, despite GenAI's visibility, only two industries (Tech and Media) show clear signs of structural disruption, while seven others remain on the wrong side of transformation.

Despite high-profile investment, industry-level transformation remains limited. GenAI has been embedded in support, content creation, and analytics use cases, but few industries show the deep structural shifts associated with past general-purpose technologies such as new market leaders, disrupted business models, or measurable changes in customer behavior.

To better quantify the state of disruption, we developed a composite AI Market Disruption Index. Each industry was scored from 0 to 5 based on five observable indicators. These scores represent normalized averages across five dimensions, triangulated from public indicators and interview-derived assessments. Alternative weighting schemes were tested to confirm consistency of industry rankings:

1. Market share volatility among top incumbents (2022 to 2025)

2. Revenue growth of AI-native firms founded after 2020
3. Emergence of new AI-driven business models
4. Changes in user behavior attributable to GenAI
5. Frequency of executive org changes attributed to AI tooling

Exhibit: GenAI disruption varies sharply by industry

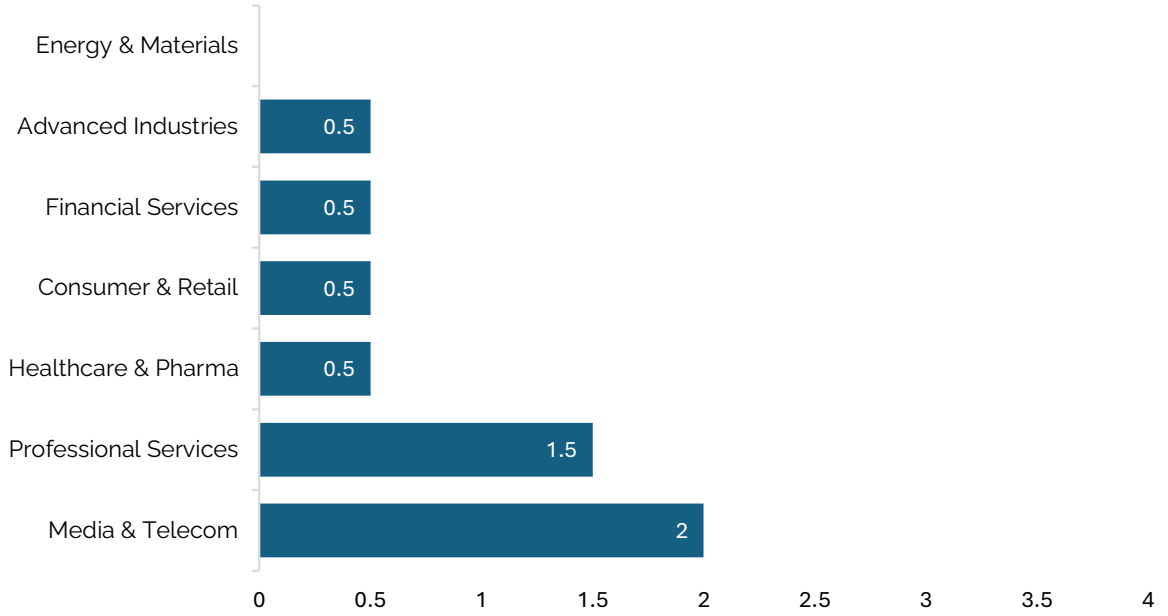


Exhibit: Description of GenAI disruption

Industry	Key Signals
Technology	New challengers gaining ground (e.g., Cursor vs Copilot); shifts in workflows
Media & Telecom	Rise of AI-native content; shifting ad dynamics; incumbents still growing
Professional Services	Efficiency gains; client delivery remains largely unchanged
Healthcare & Pharma	Documentation/transcription pilots; clinical models unchanged
Consumer & Retail	Support automation; limited impact on loyalty or leaders
Financial Services	Backend automation; customer relationships stable
Advanced Industries	Maintenance pilots; no major supply chain shifts
Energy & Materials	Near-zero adoption; minimal experimentation

Sensitivity Analysis: We tested alternative weightings for the five disruption indicators. Technology and Media & Telecom maintained top rankings across all reasonable weighting schemes, while Healthcare and Energy remained consistently low. Professional Services showed the most sensitivity to weighting changes, ranging from 1.2 to 2.1 depending on emphasis placed on efficiency gains versus structural change.

Seven out of nine major sectors showed significant pilot activity but little to no structural change. This gap between investment and disruption directly demonstrates the GenAI Divide at scale, widespread experimentation without transformation.

Interviewees were blunt in their assessments. One mid-market manufacturing COO summarized the prevailing sentiment:

"The hype on LinkedIn says everything has changed, but in our operations, nothing fundamental has shifted. We're processing some contracts faster, but that's all that has changed."

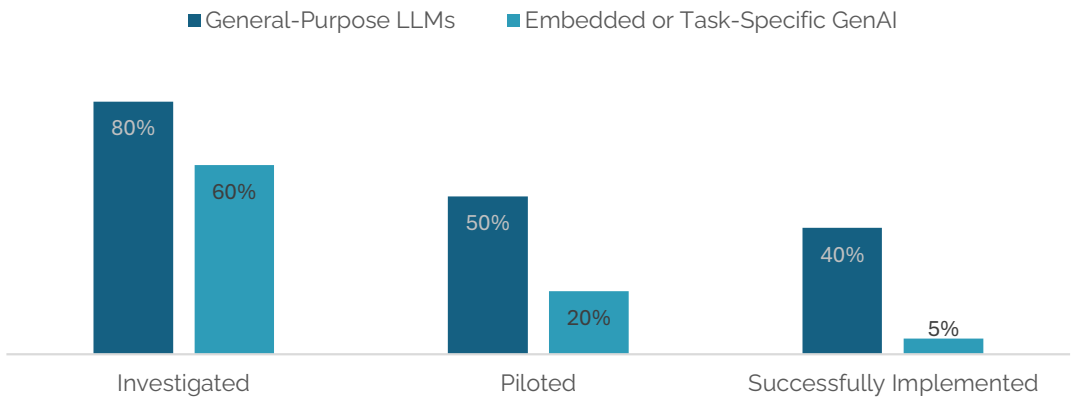
3.2 The Pilot-to-Production Chasm

Takeaway: The GenAI Divide is starkest in deployment rates, only 5% of custom enterprise AI tools reach production. Chatbots succeed because they're easy to try and flexible, but fail in critical workflows due to lack of memory and customization. This fundamental gap explains why most organizations remain on the wrong side of the divide.

Our research reveals a steep drop-off between investigations of GenAI adoption tools and pilots and actual implementations, with significant variation between generic and custom solutions.

Research Limitations: These figures are directionally accurate based on individual interviews rather than official company reporting. Sample sizes vary by category, and success definitions may differ across organizations.

Exhibit: The steep drop from pilots to production for task-specific GenAI tools reveals the GenAI divide



Research Note: We define successfully implemented for task-specific GenAI tools as ones users or executives have remarked as causing a marked and sustained productivity and/or P&L impact

The 95% failure rate for enterprise AI solutions represents the clearest manifestation of the GenAI Divide. Organizations stuck on the wrong side continue investing in static tools that can't adapt to their workflows, while those crossing the divide focus on learning-capable systems.

Generic LLM chatbots appear to show high pilot-to-implementation rates (~83%). However, this masks a deeper split in perceived value and reveals why most organizations remain trapped on the wrong side of the divide.

In interviews, enterprise users reported consistently positive experiences with consumer-grade tools like ChatGPT and Copilot. These systems were praised for flexibility, familiarity, and immediate utility. Yet the same users were overwhelmingly skeptical of custom or vendor-pitched AI tools, describing them as brittle, overengineered, or misaligned with actual workflows.

As one CIO put it, "We've seen dozens of demos this year. Maybe one or two are genuinely useful. The rest are wrappers or science projects."

While enthusiasm and budgets are often sufficient to launch pilots, converting these into workflow-integrated systems with persistent value remains rare, a pattern that defines the experience of organizations on the wrong side of the GenAI Divide.

Enterprises, defined here as firms with over $100 million in annual revenue, lead in pilot count and allocate more staff to AI-related initiatives. Yet this intensity has not translated into success. These organizations report the lowest rates of pilot-to-scale conversion.

By contrast, mid-market companies moved faster and more decisively. Top performers reported average timelines of 90 days from pilot to full implementation. Enterprises, by comparison, took nine months or longer.

Five Myths About GenAI in the Enterprise

1. *AI Will Replace Most Jobs in the Next Few Years* → Research found limited layoffs from GenAI, and only in industries that are already affected significantly by AI. There is no consensus among executives as to hiring levels over the next 3-5 years.
2. *Generative AI is Transforming Business* → Adoption is high, but transformation is rare. Only 5% of enterprises have AI tools integrated in workflows at scale and 7 of 9 sectors show no real structural change.
3. *Enterprises are slow in adopting new tech* → Enterprises are extremely eager to adopt AI and 90% have seriously explored buying an AI solution.

4. *The biggest thing holding back AI is model quality, legal, data, risk* → What's really holding it back is that most AI tools don't learn and don't integrate well into workflows.
5. *The best enterprises are building their own tools* → Internal builds fail twice as often.

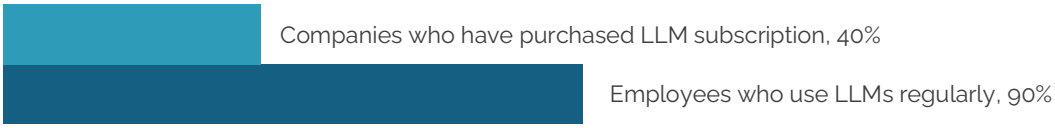
3.3 The Shadow AI Economy: A Bridge Across the Divide

Takeaway: While official enterprise initiatives remain stuck on the wrong side of the GenAI Divide, employees are already crossing it through personal AI tools. This "shadow AI" often delivers better ROI than formal initiatives and reveals what actually works for bridging the divide.

Behind the disappointing enterprise deployment numbers lies a surprising reality: AI is already transforming work, just not through official channels. Our research uncovered a thriving "shadow AI economy" where employees use personal ChatGPT accounts, Claude subscriptions, and other consumer tools to automate significant portions of their jobs, often without IT knowledge or approval.

The scale is remarkable. While only 40% of companies say they purchased an official LLM subscription, workers from over 90% of the companies we surveyed reported regular use of personal AI tools for work tasks. In fact, almost every single person used an LLM in some form for their work.

Exhibit: the shadow AI economy, employee usage far outpaces official adoption



In many cases, shadow AI users reported using LLMs multiples times a day every day of their weekly workload through personal tools, while their companies' official AI initiatives remained stalled in pilot phase.

This shadow economy demonstrates that individuals can successfully cross the GenAI Divide when given access to flexible, responsive tools. The organizations that recognize this pattern and build on it represent the future of enterprise AI adoption.

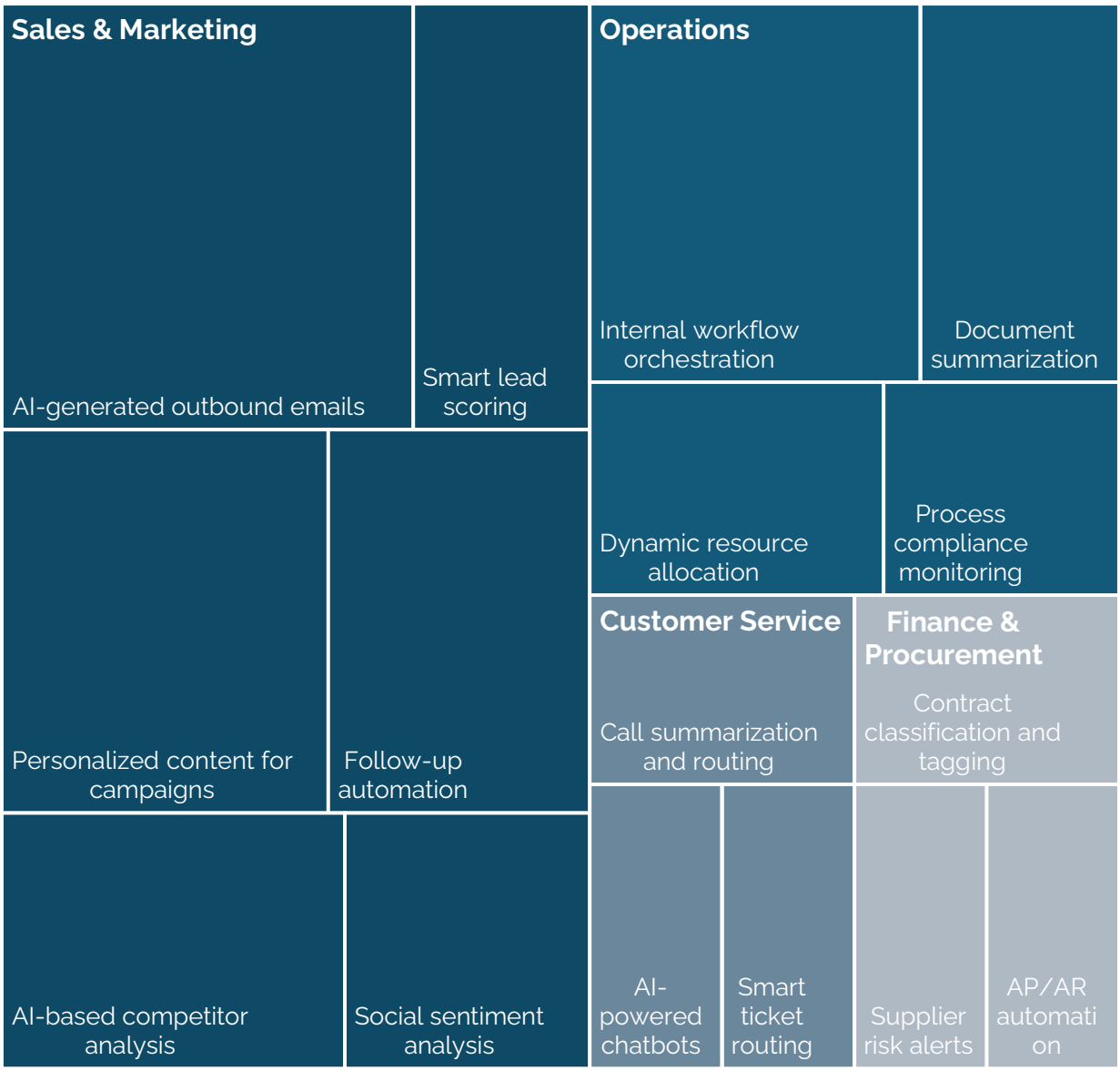
Forward-thinking organizations are beginning to bridge this gap by learning from shadow usage and analyzing which personal tools deliver value before procuring enterprise alternatives.

3.4 Investment Patterns Reflect the Divide

Takeaway: Investment allocation reveals the GenAI Divide in action, 50% of GenAI budgets go to sales and marketing, but back-office automation often yields better ROI. This bias reflects easier metric attribution, not actual value, and keeps organizations focused on the wrong priorities.

In terms of functional focus, investment in GenAI tools is heavily concentrated. As GenAI spend is not yet formally quantified across organizations, we asked executives to allocate a hypothetical $100 to different functions. Sales and marketing functions captured approximately 70 percent of AI budget allocation across organizations in our survey.

Exhibit: GenAI Investment Distribution by Function



Research Notes: While the general functional allocation of GenAI investment (e.g., ~50% to Sales & Marketing) was relatively consistent across executive interviews, the sub-category and use-case breakdowns should be treated as directional at best. Subcategories reflect synthesized notes and anecdotal patterns, rather than precise accounting. Company type drives significant variation. For example, manufacturers and healthcare providers typically directed minimal investment to Sales & Marketing and over-indexed on Operations. Tech and media firms often prioritized Marketing, Content, and Developer Productivity. Professional services leaned toward Document Automation and Legal/Compliance tools.

Sales and marketing dominate not only because of visibility, but because outcomes can be measured easily. Metrics such as demo volume or email response time align directly with board-level KPIs.

Legal, procurement, and finance functions, in contrast, offer more subtle efficiencies. These include fewer compliance violations, streamlined workflows, or accelerated month-end processes, important but hard to surface in executive conversations or investor updates.

A VP of Procurement at a Fortune 1000 pharmaceutical company expressed this challenge clearly:

"If I buy a tool to help my team work faster, how do I quantify that impact? How do I justify it to my CEO when it won't directly move revenue or decrease measurable costs? I could argue it helps our scientists get their tools faster, but that's several degrees removed from bottom-line impact."

This investment bias perpetuates the GenAI Divide by directing resources toward visible but often less transformative use cases, while the highest-ROI opportunities in back-office functions remain underfunded.

Beyond measurement challenges, trust and social proof remain decisive in purchase decisions. A Head of Procurement at a major CPG firm captured the dilemma many buyers face:

"I receive numerous emails daily claiming to offer the best GenAI solution. Some have impressive demos, but establishing trust is the real challenge. With so many options flooding our inbox, we rely heavily on peer recommendations and referrals from our network."

This highlights a broader pattern: product quality alone is rarely sufficient. Referrals, prior relationships, and VC introductions remain stronger predictors of enterprise adoption than functionality or feature set.

4 Why Pilots Stall: The Learning Gap Behind the Divide

The primary factor keeping organizations on the wrong side of the GenAI Divide is the learning gap, tools that don't learn, integrate poorly, or match workflows. Users prefer ChatGPT for simple tasks, but abandon it for mission-critical work due to its lack of memory. What's missing is systems that adapt, remember, and evolve, capabilities that define the difference between the two sides of the divide.

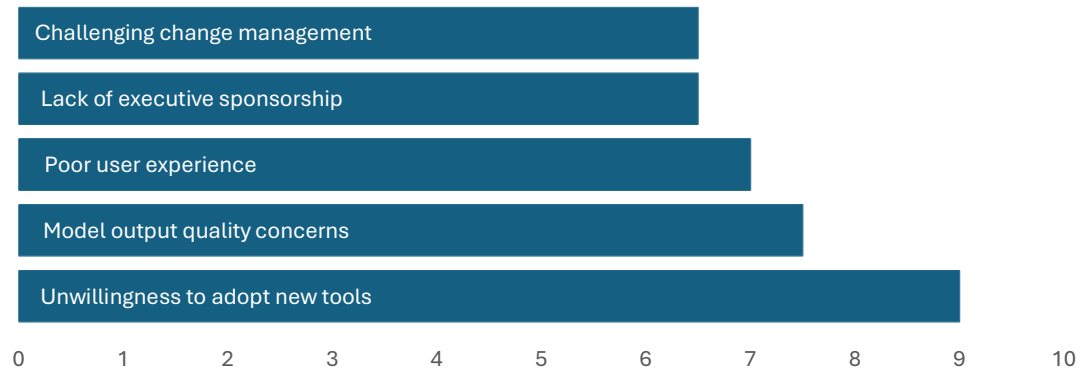
4.1 The Barriers Keeping Organizations Trapped

Takeaway: The top barriers reflect the fundamental learning gap that defines the GenAI Divide: users resist tools that don't adapt, model quality fails without context, and UX suffers when systems can't remember. Even avid ChatGPT users distrust internal GenAI tools that don't match their expectations.

To understand why so few GenAI pilots progress beyond the experimental phase, we surveyed both executive sponsors and frontline users across 52 organizations. Participants were asked to rate common barriers to scale on a 1–10 frequency scale, where 10 represented the most frequently encountered obstacles.

The results revealed a predictable leader: resistance to adopting new tools. However, the second-highest barrier proved more significant than anticipated.

Exhibit: Why GenAI pilots fail: top barriers to scaling AI in the enterprise
Users were asked to rate each issue on a scale of 1-10



Research Note: These scores reflect reported frequency rather than objective measurement of barrier impact, and may vary significantly by industry and organization size.

The prominence of model quality concerns initially appeared counterintuitive. Consumer adoption of ChatGPT and similar tools has surged, with over 40% of knowledge workers using AI tools personally. Yet the same users who integrate these tools into personal workflows describe them as unreliable when encountered within enterprise systems. This paradox illustrates the GenAI Divide at the user level.

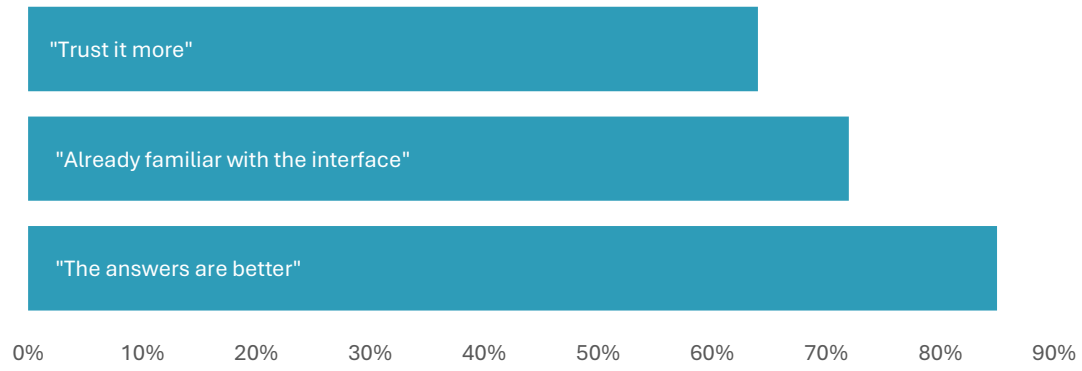
This preference reveals a fundamental tension. The same professionals using ChatGPT daily for personal tasks demand learning and memory capabilities for enterprise work. A significant number of workers already use AI tools privately, reporting productivity gains, while their companies' formal AI initiatives stall. This shadow usage creates a feedback loop: employees know what good AI feels like, making them less tolerant of static enterprise tools.

4.2 Why Generic Tools Win, And Lose

Takeaway: The GenAI Divide manifests in user preferences: ChatGPT beats enterprise tools because it's better, faster, and more familiar, even when both use similar models. But this same preference exposes why organizations remain stuck on the wrong side of the divide.

Our follow-up interviews revealed a striking contradiction. The professionals expressing skepticism about enterprise AI tools were often heavy users of consumer LLM interfaces. When asked to compare their experiences, three consistent themes emerged.

User Preference Drivers: Generic LLM Interface vs. Integrated Tool



A corporate lawyer at a mid-sized firm exemplified this dynamic. Her organization invested $50,000 in a specialized contract analysis tool, yet she consistently defaulted to ChatGPT for drafting work:

"Our purchased AI tool provided rigid summaries with limited customization options. With ChatGPT, I can guide the conversation and iterate until I get exactly what I need. The fundamental quality difference is noticeable, ChatGPT consistently produces better outputs, even though our vendor claims to use the same underlying technology."

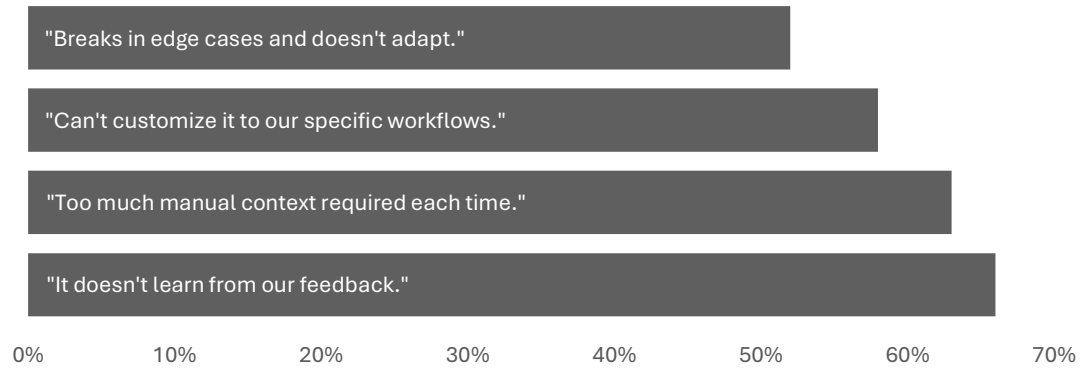
This pattern suggests that a $20-per-month general-purpose tool often outperforms bespoke enterprise systems costing orders of magnitude more, at least in terms of immediate usability and user satisfaction. This paradox exemplifies why most organizations remain on the wrong side of the GenAI Divide.

4.3 The Learning Gap That Defines the Divide

Takeaway: ChatGPT's very limitations reveal the core issue behind the GenAI Divide: it forgets context, doesn't learn, and can't evolve. For mission-critical work, 90% of users prefer humans. The gap is structural, GenAI lacks memory and adaptability.

Given users' preference for consumer LLM interfaces, we investigated what prevents broader adoption for mission-critical work. The barriers here proved distinct from general usability concerns and directly illuminated the learning gap that defines the GenAI Divide.

Exhibit: Barriers to core workflow integration



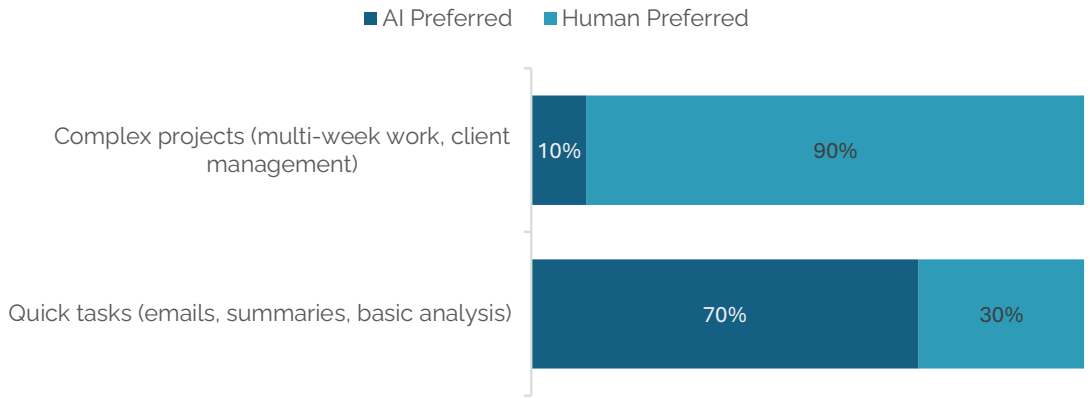
The same lawyer who favored ChatGPT for initial drafts drew a clear line at sensitive contracts:

"It's excellent for brainstorming and first drafts, but it doesn't retain knowledge of client preferences or learn from previous edits. It repeats the same mistakes and requires extensive context input for each session. For high-stakes work, I need a system that accumulates knowledge and improves over time."

This feedback points to the fundamental learning gap that keeps organizations on the wrong side of the GenAI Divide. Users appreciate the flexibility and responsiveness of consumer LLM interfaces but require the persistence and contextual awareness that current tools cannot provide.

When we asked enterprise users to rate different options for high-stakes work, the preference hierarchy became clear:

Exhibit: Perceived Fitness for High-Stakes Work
"Would you assign this task to AI or a junior colleague?"



The results reveal that AI has already won the war for simple work, 70% prefer AI for drafting emails, 65% for basic analysis. But for anything complex or long-term, humans dominate by 9-to-1 margins. The dividing line isn't intelligence, it's memory, adaptability, and learning capability, the exact characteristics that separate the two sides of the GenAI Divide.

Agentic AI, the class of systems that embeds persistent memory and iterative learning by design, directly addresses the learning gap that defines the GenAI Divide. Unlike current systems that require full context each time, agentic systems maintain persistent memory, learn from interactions, and can autonomously orchestrate complex workflows. Early enterprise experiments with customer service agents that handle complete inquiries end-to-end, financial processing agents that monitor and approve routine transactions, and sales pipeline agents that track engagement across channels demonstrate how autonomy and memory address the core gaps enterprises identify.

Exhibit: Positioning GenAI tools by customization and learning capability

	Low Memory / Learning	**High Memory / Learning**
Low Customization	Copilot, GPT wrappers	ChatGPT w/ memory (beta)
High Customization	Internal builds (fragile)	Agentic workflows, vertical SaaS

5 CROSSING THE GENAI DIVIDE: HOW THE BEST BUILDERS SUCCEED

Organizations on the right side of the GenAI Divide share a common approach: they build adaptive, embedded systems that learn from feedback. The best startups crossing the divide focus on narrow but high-value use cases, integrate deeply into workflows, and scale through continuous learning rather than broad feature sets. Domain fluency and workflow integration matter more than flashy UX.

Across our interviews, we observed a growing divergence among GenAI startups. Some are struggling with outdated SaaS playbooks and remain trapped on the wrong side of the divide, while others are capturing enterprise attention through aggressive customization and alignment with real business pain points.

The appetite for GenAI tools remains high. Several startups reported signing pilots within days and reaching seven-figure revenue run rates shortly thereafter. The standout performers are not those building general-purpose tools, but those embedding themselves inside workflows, adapting to context, and scaling from narrow but high-value footholds.

Our data reveals a clear pattern: the organizations and vendors succeeding are those aggressively solving for learning, memory, and workflow adaptation, while those failing are either building generic tools or trying to develop capabilities internally.

Winning startups build systems that learn from feedback (66% of executives want this), retain context (63% demand this), and customize deeply to specific workflows. They start at workflow edges with significant customization, then scale into core processes.

5.1 What Enterprises Actually Want: The Bridge Across the Divide

The most successful vendors understand that crossing the GenAI Divide requires building systems that executives repeatedly emphasized: AI systems that do not just generate content, but learn and improve within their environment.

When evaluating AI tools, buyers consistently emphasized a specific set of priorities. We coded these themes across interviews to quantify how often they surfaced in procurement decisions:

Exhibit: How executives select GenAI vendors
Derived from interviews and coded by category

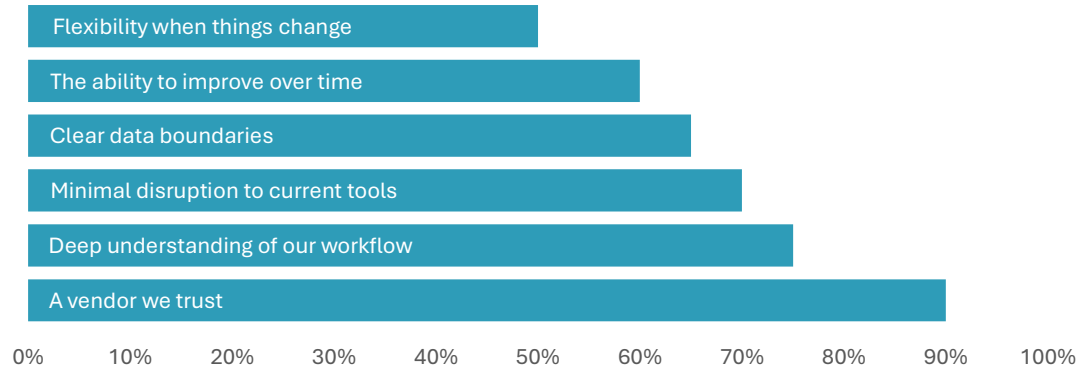


Exhibit: Direct quotes on executives selecting GenAI vendors
Derived from interviews and coded by category

What They Want	Direct Quotes
A vendor we (they) trust	"We're more likely to wait for our existing partner to add AI than gamble on a startup."
Deep understanding of our workflow	"Most vendors don't get how our approvals or data flows work"
Minimal disruption to current tools	"If it doesn't plug into Salesforce or our internal systems, no one's going to use it."
Clear data boundaries	"I can't risk client data mixing with someone else's model, even if the vendor says it's fine."
The ability to improve over time	"It's useful the first week, but then it just repeats the same mistakes. Why would I use that?"

Flexibility when things change	"Our process evolves every quarter. If the AI can't adapt, we're back to spreadsheets."

Concerns about workforce impact were far less common than anticipated. Most users welcomed automation, especially for tedious, manual tasks, as long as data remained secure and outcomes were measurable.

Despite conventional wisdom that enterprises resist training AI systems, most teams in our interviews expressed willingness to do so, provided the benefits were clear and guardrails were in place.

Despite interest in AI, there is notable skepticism toward emerging vendors, especially in high-trust or regulated workflows. Many procurement leaders told us they ignore most startup pitches, regardless of innovation.

"We receive dozens of pitches daily about AI-powered procurement tools. However, our established BPO partner already understands our policies and processes. We're more likely to wait for their AI-enhanced version than switch to an unknown vendor." , Head of Procurement, Global CPG

5.2 The Winning Playbook for Crossing the Divide

Takeaway: Startups that successfully cross the GenAI Divide land small, visible wins in narrow workflows, then expand. Tools with low setup burden and fast time-to-value outperform heavy enterprise builds. Channel referrals and peer trust are key growth levers for bridging the divide.

The most successful startups addressed both the desire for learning systems and the skepticism around new tools by executing two strategies:

5.2.1 Customizing for specific workflows

Embedding in non-critical or adjacent processes with significant customization, demonstrating clear value, then scaling into core workflows is critical. Tools that succeeded shared two traits: low configuration burden and immediate, visible value. In contrast, tools requiring extensive enterprise customization often stalled at pilot stage.

Successful categories from our sample included:

- Voice AI for call summarization and routing
- Document automation for contracts and forms
- Code generation for repetitive engineering tasks

Struggling categories were often those involving complex internal logic, opaque decision support, or optimization based on proprietary heuristics. These tools frequently hit adoption friction due to deep enterprise specificity.

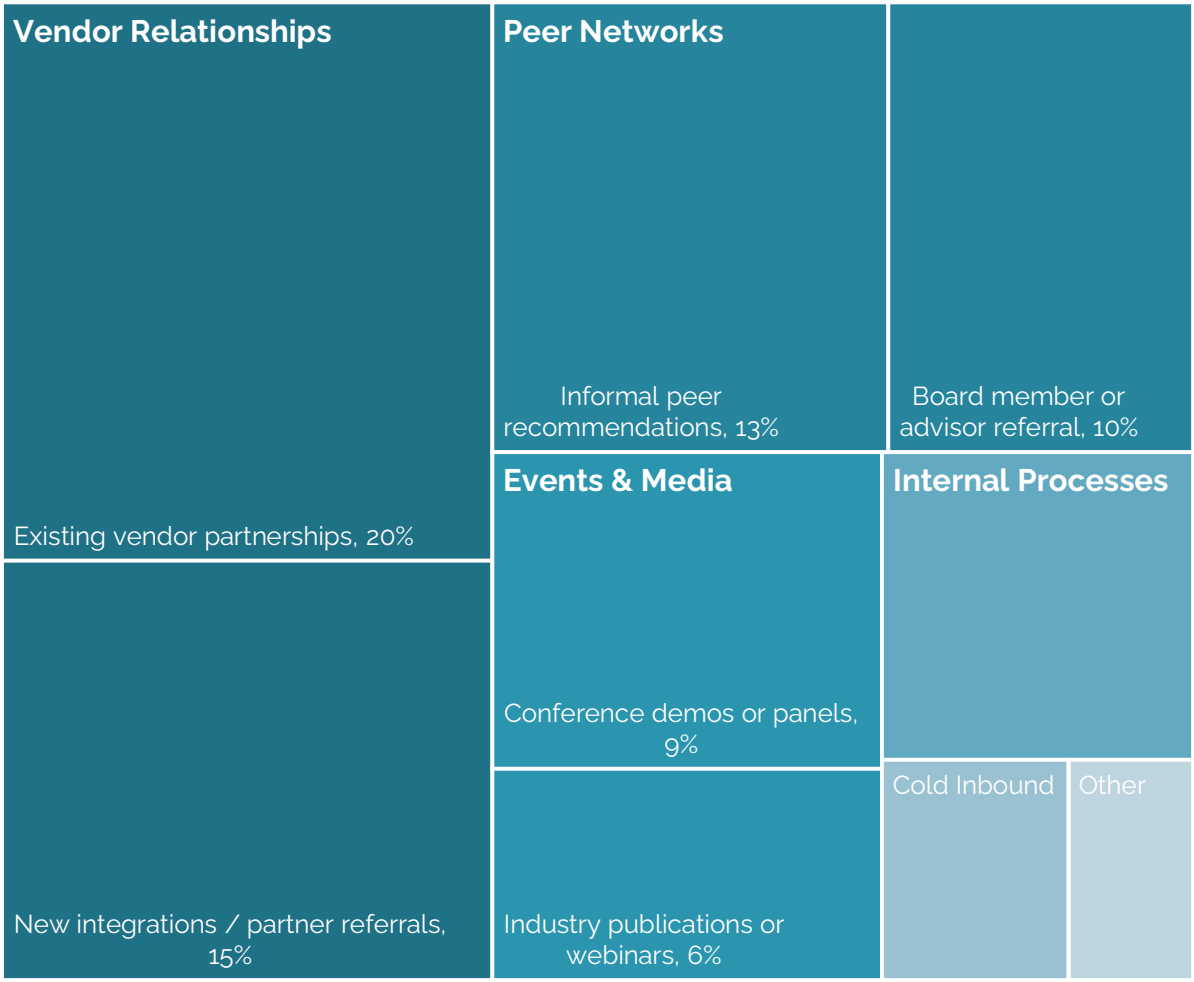
Some startups have excelled by dominating small but critical workflows, especially in sales and marketing, and then expanding. Top-quartile GenAI startups are reaching $1.2M in annualized revenue within 6–12 months of launch.

	Narrow Scope	Broad Scope
Simple Execution	Fast wins (Spend Categorization, Contract Review)	Partial pilots (Supplier Risk Monitoring)
Complex Execution	Early pilots (Negotiation Bots)	Fails (Full Procurement Orchestration)

5.2.2 Leveraging referral networks

To overcome trust barriers, successful startups often used channel partnerships with system integrators, procurement referrals from board members or advisors, and distribution through familiar enterprise marketplaces.

Exhibit: How leaders discover GenAI solutions



5.3 The Narrowing Window to Cross the Divide

Takeaway: The window for crossing the GenAI Divide is rapidly closing. Enterprises are locking in learning-capable tools. Agentic AI and memory frameworks (like NANDA and MCP) will define which vendors help organizations cross the divide versus remain trapped on the wrong side.

Enterprises are increasingly demanding systems that adapt over time. Microsoft 365 Copilot and Dynamics 365 are incorporating persistent memory and feedback loops. OpenAI's ChatGPT memory beta signals similar expectations in general-purpose tools.

Startups that act quickly to close this gap, by building adaptive agents that learn from feedback, usage, and outcomes, can establish durable product moats through both data and integration depth. The window to do this is narrow. In many verticals, pilots are already underway.

The infrastructure to support this transition is emerging through frameworks like Model Context Protocol (MCP), Agent-to-Agent (A2A), and NANDA, which enable agent interoperability and coordination. These protocols create market competition and cost efficiencies by allowing specialized agents to work together rather than requiring monolithic systems. And these frameworks form the foundation of the emerging Agentic Web, a mesh of interoperable agents and protocols that replaces monolithic applications with dynamic coordination layers.

In the next few quarters, several enterprises will lock in vendor relationships that will be nearly impossible to unwind. This 18-month horizon reflects consensus from seventeen procurement and IT sourcing leaders we interviewed, supported by analysis of public procurement disclosures showing enterprise RFP-to-implementation cycles ranging from two to eighteen months. Organizations investing in AI systems that learn from their data, workflows, and feedback are creating switching costs that compound monthly.

"We're currently evaluating five different GenAI solutions, but whichever system best learns and adapts to our specific processes will ultimately win our business. Once we've invested time in training a system to understand our workflows, the switching costs become prohibitive." - CIO, $5B Financial Services Firm

6 Crossing the GenAI Divide: How the Best Buyers Succeed

Organizations that successfully cross the GenAI Divide approach AI procurement differently, they act like BPO clients, not SaaS customers. They demand deep customization, drive adoption from the front lines, and hold vendors accountable to business metrics. The most successful buyers understand that crossing the divide requires partnership, not just purchase.

Across our interviews, one insight was clear: the most effective AI-buying organizations no longer wait for perfect use cases or central approval. Instead, they drive adoption through

distributed experimentation, vendor partnerships, and clear accountability. These buyers are not just more eager, they are more strategically adaptive.

In our sample, external partnerships with learning-capable, customized tools reached deployment ~67% of the time, compared to ~33% for internally built tools. While these figures reflect self-reported outcomes and may not account for all confounding variables, the magnitude of difference was consistent across interviewees.

This gap explains why ChatGPT dominates for ad-hoc tasks but fails at critical workflows, and why generic enterprise tools lose to both consumer LLMs and deeply customized alternatives.

6.1 Organizational Design for Crossing the Divide

Takeaway: The right organizational structure is critical for crossing the GenAI Divide. Strategic partnerships are twice as likely to succeed as internal builds. Success depends less on resources and more on decentralizing authority with clear ownership.

The dominant barrier to crossing the GenAI Divide is not integration or budget, it is organizational design. Our data shows that companies succeed when they decentralize implementation authority but retain accountability.

We observed three primary team structures for GenAI implementation, with materially different outcomes that reflect which side of the divide organizations land on.

	% of Deployments	
Strategic Partnerships *(Buy)*	**66%**	Procure external tools, co-develop with vendors
Internal Development *(Build)*	**33%**	Build and maintain GenAI tools fully in-house
Hybrid *(Build-Buy)*	*Insufficient data to quantify*	Internal team co-develops with an external vendor

Research Limitations: *These percentages reflect our interview sample of 52 organizations and may not represent broader market patterns. Success definitions varied across organizations, and observation periods may not capture long-term implementation trends.*

Important Limitation: *These success rate differences may reflect organizational capabilities rather than implementation approach alone. Organizations choosing external partnerships may have different risk tolerance, procurement sophistication, or internal technical capacity than those building internally. The correlation between external partnerships and success does not necessarily prove causation.*

Strategic partnerships achieved a significantly higher share of successful deployments than internal development efforts. While we observed far more BUILD initiatives than BUY initiatives in our sample, with many more organizations exploring internal development, the

success rates favored external partnerships. Though we lack precise data on total initiative volumes, the pattern suggests that internal development efforts have substantially lower success rates despite being more commonly attempted.

While organizations often combined approaches, pilots built via strategic partnerships were 2x as likely to reach full deployment as those built internally. More strikingly, employee usage rates were nearly double for externally built tools.

These partnerships often provided faster time-to-value, lower total cost, and better alignment with operational workflows. Companies avoided the overhead of building from scratch, while still achieving tailored solutions. Organizations that understand this pattern position themselves to cross the GenAI Divide more effectively.

6.2 Buyer Practices that Cross the Divide

Across interviews, a consistent pattern emerged among organizations successfully crossing the GenAI Divide: top buyers treated AI startups less like software vendors and more like business service providers, holding them to benchmarks closer to those used for consulting firms or BPOs. These organizations:

- Demanded deep customization aligned to internal processes and data
- Benchmarked tools on operational outcomes, not model benchmarks
- Partnered through early-stage failures, treating deployment as co-evolution
- Sourced AI initiatives from frontline managers, not central labs

In this last pattern, individual contributors and team managers often played a critical role. Many of the strongest enterprise deployments began with power users, employees who had already experimented with tools like ChatGPT or Claude for personal productivity. These "prosumers" intuitively understood GenAI's capabilities and limits, and became early champions of internally sanctioned solutions. Rather than relying on a centralized AI function to identify use cases, successful organizations allowed budget holders and domain managers to surface problems, vet tools, and lead rollouts. This bottom-up sourcing, paired with executive accountability, accelerated adoption while preserving operational fit.

6.3 Where the Real ROI Lives: Beyond the Divide

Takeaway: Organizations that cross the GenAI Divide discover that ROI is often highest in ignored functions like operations and finance. Real gains come from replacing BPOs and external agencies, not cutting internal staff. Front-office tools get attention, but back-office tools deliver savings.

Despite 50% of AI budgets flowing to sales and marketing (from the theoretical estimate with executives), some of the most dramatic cost savings we documented came from back-office automation. While front-office gains are visible and board-friendly, the back-office deployments often delivered faster payback periods and clearer cost reductions.

Best-in-class organizations are generating measurable value across both areas:

Front-office wins:

- Lead qualification speed: 40% faster
- Customer retention: 10% improvement through AI-powered follow-ups and messaging

Back-office wins:

- BPO elimination: $2-10M annually in customer service and document processing
- Agency spend reduction: 30% decrease in external creative and content costs
- Risk checks for financial services: $1M saved annually on outsourced risk management

Notably, these gains came without material workforce reduction. Tools accelerated work, but did not change team structures or budgets. Instead, ROI emerged from reduced external spend, eliminating BPO contracts, cutting agency fees, and replacing expensive consultants with AI-powered internal capabilities.

The pattern suggests that while sales and marketing capture the majority of attention and investment, back-office automation may offer more dramatic and sustainable returns for organizations willing to look beyond the obvious use cases and truly cross the GenAI Divide.

6.4 The Job Impact Reality: What Crossing the Divide Actually Means

Takeaway: GenAI is already starting to have workforce impact and it is manifesting through selective displacement of previously outsourced functions and constrained hiring patterns, but not through broad-based layoffs. Organizations that have crossed the GenAI Divide demonstrate measurable external cost reduction while slightly decreasing internal headcount.

6.4.1 Displacement Patterns and Organizational Strategy

Our analysis reveals that GenAI-driven workforce reductions concentrate in functions historically treated as non-core business activities: customer support operations, administrative processing, and standardized development tasks. These roles exhibited vulnerability prior to AI implementation due to their outsourced status and process standardization. Executives were hesitant to reveal the scope of layoffs due to AI but it was between 5-20% of customer support operations and administrative processing work in these companies.

Industry-specific hiring expectations reveal a clear correlation with GenAI impact patterns. In sectors showing minimal structural disruption from AI, Healthcare, Energy, Advanced Industries, most executives report no current or anticipated hiring reductions over the next five years. A few executives mentioned that they could anticipate decreased hiring but admitted that they currently do not have the systems in place to accurately predict when or where it could happen. Healthcare executives, for example, express no expectation of reducing physician or clinical staff hiring.

Conversely, in Technology and Media sectors where GenAI has demonstrated measurable impact, >80% of executives anticipate reduced hiring volumes within 24 months.

Critical finding**:** This dynamic remains concentrated among advanced AI adopters, and only in industries that are currently experiencing significant disruption with GenAI (tech, media).

6.4.2 Evolving Hiring Criteria and Skills Requirements

GenAI adoption creates divergent hiring strategies across organizations. While executives demonstrate no consensus regarding entry-level or general hiring volumes, they consistently emphasize AI literacy as a fundamental capability requirement. This reflects organizational recognition that AI proficiency represents a competitive advantage in workflow optimization.

"Our hiring strategy prioritizes candidates who demonstrate AI tool proficiency. Recent graduates often exceed experienced professionals in this capability." - VP of Operations, Mid-Market Manufacturing

6.4.3 Future Workforce Impact Projections

MIT's Project Iceberg analysis provides quantitative context for potential automation exposure:

Current automation potential: 2.27% of U.S. labor value

Latent automation exposure: $2.3 trillion in labor value affecting 39 million positions

This latent exposure becomes actionable as AI systems develop persistent memory, continuous learning, and autonomous tool integration, capabilities that define crossing the GenAI Divide.

Workforce transformation will occur gradually rather than through discrete displacement events. Until AI systems achieve contextual adaptation and autonomous operation, organizational impact will manifest through external cost optimization rather than internal restructuring.

6.5 Beyond Agents: The Agentic Web

Takeaway: The next evolution beyond individual AI agents is an agentic web where autonomous systems can discover, negotiate, and coordinate across the entire internet infrastructure, fundamentally changing how business processes operate.

The infrastructure foundations for this transformation are already emerging through protocols like Model Context Protocol (MCP), Agent-to-Agent (A2A), and NANDA, which enable not just agent interoperability but autonomous web navigation. In an agentic web, systems will autonomously discover optimal vendors and evaluate solutions without human research, establish dynamic API integrations in real-time without pre-built connectors, execute trustless transactions through blockchain-enabled smart contracts, and develop emergent workflows that self-optimize across multiple platforms and organizational boundaries. Early experiments show procurement agents identifying new suppliers and negotiating terms independently, customer service systems coordinating seamlessly across platforms, and content creation workflows spanning multiple providers with automated quality assurance

and payment. This represents a fundamental shift from today's human-mediated business processes to autonomous systems that operate across the entire internet ecosystem, moving well beyond the current GenAI Divide to reshape how organizations discover, integrate, and transact in a networked economy.

7 Conclusion: Bridging the GenAI Divide

Organizations that successfully cross the GenAI Divide do three things differently: they buy rather than build, empower line managers rather than central labs, and select tools that integrate deeply while adapting over time. The most forward-thinking organizations are already experimenting with agentic systems that can learn, remember, and act autonomously within defined parameters.

This transition marks not just a shift in tooling, but the emergence of an Agentic Web: a persistent, interconnected layer of learning systems that collaborate across vendors, domains, and interfaces. Where today's enterprise stack is defined by siloed SaaS tools and static workflows, the Agentic Web replaces these with dynamic agents capable of negotiating tasks, sharing context, and coordinating action across the enterprise.

Just as the original Web decentralized publishing and commerce, the Agentic Web decentralizes action, moving from prompts to autonomous protocol-driven coordination. Systems like NANDA, MCP, and A2A represent early infrastructure for this web, enabling organizations to compose workflows not from code, but from agent capabilities and interactions. As enterprises begin locking in vendor relationships and feedback loops through 2026, the window to cross the GenAI Divide is rapidly narrowing. The next wave of adoption will be won not by the flashiest models, but by the systems that learn and remember and/or by systems that are custom built for a specific process.

The shift from building to buying, combined with the rise of prosumer adoption and the emergence of agentic capabilities, creates unprecedented opportunities for vendors who can deliver learning-capable, deeply integrated AI systems. The organizations and vendors that recognize and act on these patterns will establish the dominant positions in the post-pilot AI economy, on the right side of the GenAI Divide.

For organizations currently trapped on the wrong side, the path forward is clear: stop investing in static tools that require constant prompting, start partnering with vendors who offer custom systems, and focus on workflow integration over flashy demos. The GenAI Divide is not permanent, but crossing it requires fundamentally different choices about technology, partnerships, and organizational design.

8 Appendix

8.1 Acknowledgments

Produced in collaboration with Project NANDA out of MIT: NANDA (Networked Agents And Decentralized Architecture) builds on Anthropic's Model Context Protocol (MCP) and the Google/Linux Foundation A2A to create infrastructure for distributed agent intelligence

at scale. Our research focuses on translating AI capabilities into measurable business outcomes across enterprise and mid-market organizations. We acknowledge the generous participation of executives who shared their implementation experiences and insights.

8.2 Research Methodology and Limitations

Methodology: 52 structured interviews across enterprise stakeholders, systematic analysis of 300+ public AI initiatives and announcements, and surveys with 153 leaders. Success defined as deployment beyond pilot phase with measurable KPIs. ROI impact measured 6 months post-pilot, adjusted for department size. Confidence intervals calculated using bootstrap resampling methods where applicable.

Sample Limitations:

- Our sample may not fully represent all enterprise segments or geographic regions
- Organizations willing to discuss AI implementation challenges may systematically differ from those declining participation, potentially creating bias toward either more experimental or more cautious adopters
- Selection bias possible in organizations willing to participate in AI research
- Success metrics vary significantly across organizations and industries, limiting direct comparisons

Methodological Constraints:

- Industry disruption scores reflect publicly observable patterns and may not capture private or emerging developments
- Build vs. buy percentages based on interview responses rather than comprehensive market data
- ROI measurements complicated by concurrent operational improvements and external economic factors
- Six-month observation period may be insufficient to fully assess "successful deployment" for complex enterprise systems, potentially understating success rates for longer-term implementations

External Factors Not Fully Addressed:

- Regulatory constraints affecting adoption

8.3 Research Instruments

This study used two semi-structured protocols:

- **Executive interviews** focused on investment decisions, organizational design, and vendor selection

- **Functional leader interviews** explored user preferences, workflow fit, and friction in daily usage

A lightweight survey supplemented these interviews with quantifiable input on tool adoption, satisfaction, and barriers.

8.3.1 Executive Interview Questionnaire

Focus: Strategy, investment, procurement, outcomes

Section 1: Strategy and Budget

1. Has your organization allocated a dedicated budget for GenAI initiatives?

2. Which business functions are currently prioritized?

3. Are there specific use cases identified for GenAI in your org?

Section 2: Buy vs Build

4. Do you primarily build internally, partner externally, or take a hybrid approach?

5. What drives that decision, cost, risk, timeline, control, etc.?

Section 3: Pilot to Scale

6. How many GenAI pilots have been launched since Jan 2024?

7. Of those, how many are now deployed at scale?

8. What were the major barriers that stalled scale-up?

Section 4: Procurement and Evaluation

9. How do you evaluate potential GenAI vendors or partners?

10. What are the most important selection criteria (e.g., trust, integration, data control)?

11. How do referrals or ecosystem partners factor into decisions?

Section 5: ROI and Outcomes

12. Have you observed measurable ROI from any GenAI deployment?

13. Which metrics (cost savings, productivity, customer retention) were used?

14. Were there specific back-office or front-office gains?

Section 6: Workforce and Governance

15. Have you reduced headcount due to GenAI?

16. Who leads implementation efforts (e.g., IT, line managers, AI CoE)?

17. How are responsibilities distributed across teams?

8.3.2 Functional Leader / User Interview Questionnaire

Focus: Tool usability, friction, and performance in day-to-day workflows

Section 1: Personal Use and Preferences

1. Do you personally use GenAI tools like ChatGPT or Claude? For what tasks?

2. Do you use internal GenAI tools at work? How do they compare?

Section 2: Enterprise Tool Experience

3. What GenAI tools have been introduced by your organization?

4. How frequently do you use them?

5. What's working well? What's frustrating?

Section 3: Workflow Fit

6. Do these tools integrate with your core systems (e.g., CRM, internal portals)?

7. Do they adapt to your workflow over time or feel static?

8. Have you seen them improve from user feedback?

Section 4: Task Type Preferences

9. For [X use case: email, doc drafting, research, etc.], would you prefer AI or a human colleague?

10. What kinds of tasks do you trust AI with? What kinds do you avoid?

Section 5: Adoption Barriers

11. What stops you or your colleagues from using these tools more often?

12. Are training, UX, or trust in outputs major issues?

[1] Project Iceberg - Are you living under the Agentic API?