

Don't Get Fooled Again: The Real Story Behind AI's 95% PROBLEM



1. Introduction

*95% of enterprise GenAI efforts show no measurable profit and loss (P&L) impact.*¹

You have probably seen that quote making the rounds—splashed across LinkedIn posts and breathless tech headlines. It is a provocative number, the kind that fuels skepticism: *Maybe all this AI talk really is hype*. And to be fair, skepticism is not misplaced. There has been plenty of AI theater—grand demos, half-baked pilots, and buzzword-laden decks that fail to deliver.

Here is what those headlines miss: AI does pay off when organizations create the conditions for it to pay off. When you dig into the research behind that viral “95%” figure— MLQ / MIT NANDA *State of AI in Business 2025* report (the NANDA paper)—you find a story not of failure, but of underdeveloped practice. Most projects are stuck in what researchers call pre-scale mode: scattered pilots, shallow integration, weak measurement, and little follow-through.

What the “95%” Statistic Does—and Does Not—Mean
What the report actually says: The MLQ / MIT NANDA <i>State of AI in Business 2025</i> report finds that roughly 95% of enterprise GenAI initiatives show no measurable profit and loss (P&L) impact under current practices.
What that number measures: The denominator includes pilots and production deployments where measurable P&L data was available. It does not include the vast number of informal or small-scale uses of AI that have not yet been linked to business metrics.
How “success” is defined: The report focuses narrowly on GenAI-focused projects where financial impact could be explicitly measured — typically through revenue growth, cost savings, or productivity gains that appear in accounting systems.
What it does <u>not</u> mean: The statistic does not imply that AI is failing as a technology. Instead, it highlights a maturity gap: most organizations have not yet developed the operational discipline, governance, or measurement frameworks needed to capture and report AI’s real value.
(Source: MLQ / MIT NANDA. “The GenAI Divide: State of AI in Business 2025.” Version 0.1, July 2025, MLQ. https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf)

¹ MLQ / MIT NANDA. *The GenAI Divide: State of AI in Business 2025*. Version 0.1, July 2025, MLQ, https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf (AKA the NANDA paper).

Meanwhile, MIT's Center for Information Systems Research (CISR) paints the other half of the picture. Their longitudinal data shows that companies that have scaled AI as a way of working—not a collection of experiments—outperform their peers in growth, profitability, and innovation. Both studies describe the same curve, just from different vantage points. The 95% have not built the conditions for AI to succeed yet. The 5% have.²

This paper is about crossing that gap. It is a practical guide for Alaska's IT leaders to move from experimentation to earnings—by building the organizational conditions that make AI value repeatable, measurable, and real.

2. Why The “95%” Statistic Is Both True and Misleading

The viral “95%” statistic makes for great clickbait. It taps into a real anxiety: *what if AI is all smoke and mirrors?* But the underlying research tells a subtler story. The NANDA paper does not claim AI is failing; it shows that most organizations are still stuck in pilot purgatory. Their efforts do not connect to operations, finance, or governance. The issue is not that AI fails; **it is that the way we work with AI does not work.**

CISR's research offers the counterpoint. They track enterprise AI maturity through four stages, each requiring distinct activities, roles, and structures to move forward:

1. **Experimentation:** This is the creative playground phase—isolated pilots and proofs of concept flourish here. Innovation teams, data scientists, or even enthusiastic departments try out models and tools to test what is possible. However, most efforts remain disconnected from core operations. To mature, organizations need to introduce lightweight governance, basic documentation standards, and clear criteria for success. Think of this phase as exploring the terrain before building roads.
2. **Scaling:** Here, the focus shifts from novelty to repeatability. Teams begin establishing shared platforms, reusable data pipelines, and model deployment frameworks. Technical excellence matters, but so does operational discipline—creating a central AI hub or Center of Excellence often helps unify these efforts. Processes like model validation, access control, and version management emerge. The transition from Experimentation to Scaling depends on executive sponsorship and cross-functional alignment.
3. **Ways of Working:** At this stage, AI becomes embedded in daily business operations. Business and technical teams collaborate through structured workflows; data engineers, analysts, and product owners work side-by-side. AI literacy programs expand beyond IT, and measurement frameworks tie outcomes directly to financial or performance metrics. Organizationally, this is where a culture of iteration and continuous improvement replaces ad hoc experimentation.
4. **Ecosystem:** The final stage represents a self-sustaining AI organization—one that learns from itself. External partnerships, open innovation, and federated governance allow AI to scale across

² Woerner, Stephanie L., Ina Sebastian, Peter Weill, and Evgeny Káganer. *Grow Enterprise AI Maturity for Bottom-Line Impact*. MIT Center for Information Systems Research, Aug. 2025, https://cistr.mit.edu/publication/2025_0801_EnterpriseAIMaturityUpdate_WoernerSebastianWeillKaganer. (AKA the CISR paper.)

business lines. Feedback loops become institutional, and the system improves over time. Moving into this stage often requires robust ethics frameworks, clear accountability models, and the ability to retrain both people and algorithms as new data and opportunities arise.

The biggest performance leap happens between Stages 2 and 3—when AI stops being a project and starts being a practice. Firms that make that shift show clear, measurable returns, not only in efficiency but in adaptability and long-term innovation capacity.

Understanding the maturity curve is essential; understanding why organizations stall on it turns potential into performance.

3. Why Most Efforts Stall Before the Payoff

Understanding these stages is useful—but it also raises a harder question: if the path is clear, why do so many organizations stall before reaching the payoff? The answer lies in what researchers call “pre-scale mode,” a pattern where promising experiments never mature into lasting capabilities.

The pre-scale pattern is characterized by:

- **One-off experiments** that never reach production. These tend to live inside innovation teams, power users, or R&D labs and never transition into core systems. Without defined owners or integration pathways, even promising prototypes get abandoned after the initial excitement fades.
- **No feedback loops or memory**—systems do not learn from real use. Most pilot projects are built as static one-offs without mechanisms for gathering user feedback, tracking outcomes, or retraining models. Without these learning loops, trust and performance stagnate.
- **Siloed data and manual inputs** that crumble when scaled. Data pipelines are often bespoke or department-specific, preventing reuse across teams. Manual data wrangling or inconsistent definitions create friction that multiplies with every new use case.
- **Governance that shows up too late**, treating risk management as an obstacle instead of a design principle. Many organizations bolt on compliance or ethical review after deployment, rather than embedding guardrails from the start. This reactive approach slows progress and erodes stakeholder confidence.
- **Unmeasured results**, which makes it impossible to prove ROI or justify continued investment. Without clear metrics or consistent measurement frameworks, leaders cannot connect AI outcomes to strategic or financial goals. This lack of evidence traps projects in perpetual pilot mode.

“When pilots live in isolation, the payoff lives there too.”

The pattern is simple: AI fails to move the needle not because it is ineffective, but because it is disconnected. When pilots live in isolation, the payoff lives there too.

4. The Hidden Value Problem

Even when organizations clear the early hurdles of adoption, another challenge emerges: recognizing the value that is already being created. Savings might show up in the wrong cost center or be delayed by procurement cycles. Improvements in cycle time, customer satisfaction, and compliance risk often go unquantified and unnoticed, masking real progress. It is a measurement problem, not a performance problem. In other words, value creation is happening—but organizations are failing to recognize and capture it.

When Getting More Done Does Not Show Up on the P&L

A growing body of research, including the NANDA report and multiple academic studies, suggests another hidden factor: individual productivity is rising thanks to tools like ChatGPT and Copilot, but organizations struggle to measure it.

Individual contributors are already drafting faster, coding faster, and summarizing faster with tools like ChatGPT and Copilot. In controlled settings, these gains are measurable; in daily operations, they dissipate unless organizations connect tools to workflows, telemetry, and baselines. Controlled studies confirm this: MIT researchers found that workers using ChatGPT completed writing tasks 11 minutes faster with 18% higher quality³; Microsoft’s field research showed developers using Copilot produced 7–20% more pull requests per week⁴.

These efficiency gains are real—but they are fragmented. Each employee is quietly getting more done, yet those improvements do not aggregate neatly into corporate performance metrics. As the NANDA paper notes, most GenAI adoption today happens at the individual contributor level, outside formal programs or pilots. That means higher throughput and lower friction in daily work, but no direct line of sight to the P&L.

Until organizations connect these personal tools to workflows, data systems, and governance structures, their impact remains diffuse—significant in practice, invisible in accounting. Solving the hidden value problem requires treating AI as part of the work system — with observable workflows, owned metrics, and explicit attribution.

5. Building the Conditions for AI to Pay Off

Pango Technology’s earlier white paper, *Adopting AI in Alaska: A Roadmap for 2025*⁵, laid out a five-stage adoption framework: Get Ready → Start Small → Find the Big Idea → Scale Safely → Continuous Learning. That model remains solid. What is new is understanding how to move from “start small” to

³ Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond. “Study Finds ChatGPT Boosts Worker Productivity in Writing Tasks.” *MIT News*, Massachusetts Institute of Technology, 14 July 2023, <https://news.mit.edu/2023/study-finds-chatgpt-boosts-worker-productivity-writing-0714>

⁴ Cui, Kevin Zheyuan, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz. “*The Productivity Effects of Generative AI: Evidence from a Field Experiment with GitHub Copilot*.” MIT GenAI, 27 Mar. 2024, <https://mit-genai.pubpub.org/pub/v5iixksv>.

⁵ Pango Technology. “*Adopting AI in Alaska: A Roadmap for 2025*.” Pango Knowledge, 27 August 2025, <https://knowledge.pangotechnology.com/adopting-ai-in-alaska-a-roadmap-for-2025/>.

“scale safely” without losing momentum—or credibility.

Focus on P&L-adjacent work—processes where measurable cost, speed, or risk outcomes already exist.

Think:

- Reducing third-party spending (e.g., agencies or contractors).
- Speeding up service cycles.
- Improving data quality or risk controls.
- Fixing leakage points in billing, reporting, or claims.

6. From Pilots to Practice

To break the “95%” cycle, organizations must move beyond ad hoc pilots and toward a repeatable process that blends creativity with control. Think of it as an industrial-strength learning loop—each step designed to preserve momentum while tightening feedback and accountability. The goal is to turn discovery into delivery without losing the spark of innovation—or overlooking real productivity and quality gains that often emerge well before they show up on the P&L.

- **Idea & Scoping:** Begin with a grounded business hypothesis, not a technology curiosity. Define clear success metrics—cost savings, efficiency gains, risk reduction—and, if possible, map exactly where they will show up in the P&L.
- **Feasibility Spike:** Develop a minimal, governed prototype that tests your riskiest assumptions. Keep it small and timeboxed—two to four weeks—so failures are inexpensive, and lessons are fast.
- **In-Workflow Pilot:** Move beyond controlled environments and embed the prototype into real workflows. Involve end users early and treat this phase as a rehearsal for scale.
- **Scale Gate:** Require verified metrics and formal risk review before rollout.
- **Sustain:** Maintain a quarterly rhythm of review and optimization. Sustaining is not maintenance—it is an engine of compounding improvement.

This funnel converts exploration into an operating capability with owners, budgets, and audited results.

7. Alaska-Specific Realities

Alaska presents unique operational challenges—and opportunities for AI to make a difference:

- **Geography:** Automation can help bridge long-distance coordination gaps.
- **Workforce:** With limited local talent pools, AI becomes a workforce multiplier.
- **Regulation:** The public sector can lead by example if risk and ethics are built in early.
- **Small markets:** Lightweight cloud AI tools make advanced capabilities accessible to small and mid-size enterprises.

As our earlier paper emphasized, AI is not about chasing hype; it is about building resilience and capacity in an environment where those qualities matter.

8. The Bottom Line

AI is not failing—our approach is. The '95%' number is not a verdict on AI. It is a map. It shows where most organizations are today and where the opportunity lies.

For Alaska's business and IT leaders, the path forward is simple but demanding. Treat AI not as a side project, but as a way of working. Build business systems that learn, recognize real productivity and quality gains—whether or not they're immediately visible in financial systems—and scale what works.

At Pango Technology, we call that "turning experiments into earnings." And we are here to help Alaskan organizations do exactly that.

Build the conditions. Measure like finance. Scale what works.