

To drive value for the enterprise, AI doesn't need to be smarter. It needs to be more savvy – and equipped to navigate the messy modern organization.

Research from Parable conducted across 1,400+ enterprise tasks empirically validates the benefits of organizational context for enterprise AI deployments. As part of its platform vision, Parable proposes future architectures to consider as organizations continue to wrestle with the most effective and accountable approaches to implementing AI and agents at scale.

Everyone agrees AI should be transforming enterprise operations – and almost no one can definitively prove that it is. Multiple studies over the last six months have documented this disconnect, termed the "AI value gap." Many companies continue to invest in AI on the basis that they'll eventually trend upwards on the 'J-curve,' but the cost outlay without accountability is becoming harder to ignore. Implementation and compute costs continue to rise, but results at scale have been largely disappointing – especially as reflected on the balance sheet. Many leaders we've spoken to cite positive, employee-level productivity gains that ultimately fail to convert to efficiencies and value for the business.

In sum, the models keep getting smarter; more employees are adopting AI tools. And yet the ROI attributed directly to AI investments and implementations remains uneven— sporadic instead of systemic, and challenging to capture quantitatively.

Earlier this year, AI research and training company [Mercor](#) released their APEX-Agents benchmark, which was covered by [TechCrunch](#) and other outlets and open-sourced on Hugging Face. They tested frontier models—including those from OpenAI, Anthropic, and Google—against 480 realistic enterprise tasks across law, investment banking, and management consulting. The best models completed roughly 25% of tasks on their first attempt: impressive relative to where AI was even last year, but by no means up to par with trained human workers (note: the current position as of late Q1'2026 is 36% of tasks completed).

So the question follows: if AI is so smart, what hurdles are the models encountering? And how can we meaningfully harness the power of AI for the enterprise?

The Counterintuitive Finding: It's Not a Reasoning Problem

At Parable, we've found that the barrier to successful, systemic agent deployment has less to do with model intelligence. In Q1, we tested our context graph methodology against Mercor's [APEX-Agents benchmark](#) to

empirically probe a belief we've held for nearly two years: modern enterprise organizations are messy, sprawling, and complex, and graphing the nodes and relationships that exist within them provides the infrastructure and intelligence layer for AI implementations that drive real operating leverage.

Context graphs have been hotly debated—and now co-opted into promotional statements and pitch decks—since Foundation Capital published their now-viral article in late 2025. Defined simply, a context graph is a structured, queryable map of how work actually happens across an organization: who is working on what, where and when decisions are made, and how information flows between teams and tools. It's assembled from the signals teams already produce in calendar invites, emails, chats, tickets, and CRM notes. Parable has been building its context graph for more than two years, and our cofounder and CEO [wrote a detailed response](#) to the context graph discussion in January.

Research Scope & Findings

The discussion surrounding context graphs has been largely theoretical versus application-driven, so we put context graphs to the test.

Using the APEX-Agents benchmark, we ran 1,400+ task evaluations across 8 experimental runs and 3 professional domains. What we found: 70% of enterprise tasks in the benchmark were fundamentally retrieval problems. Agents faltered because they couldn't find the right information in the right place—the challenge was navigational, not computational.

The hurdle the AI encountered was orientation—a problem that might sound familiar to any precocious new hire. Agents were dropped into complex organizational environments with no map and asked to perform like seasoned professionals who've spent years learning where information tends to live and where to look for context or infer association.

The data painted a clear picture: when agents searched flat across all available content, they found the right information roughly 41% of the time. When they knew which document to search, that number jumped to 91%. The entire gap is attributable to navigation—knowing where to look and how to proceed from there.

When agents had access to Parable's context graph as a queryable tool—not context dumped into a prompt, but a structured resource they could consult on demand—performance noticeably improved:

- They **complete tasks nearly 2x faster**, with **64% fewer computational resources** consumed
- **+10 percentage points in law**, where document navigation is a primary challenge and nuance is common
- **+3 percentage points overall** across 361 paired evaluations (statistically significant)

The goal of this research was never to make the model marginally smarter. What agents needed to perform more effectively—and more efficiently—was organizational awareness.

What Worked, What Didn't—And Why It Matters

Our research showed that context graphs create the most value in complex, cross-functional workflows where information is scattered and tribal knowledge is the primary bottleneck.

They help least when a given task is pure computation, or when an agent can already navigate well on its own.

This maps directly to enterprise operational reality: some workflows are well-suited to robotic process automation, whereas others will require context-augmented orchestration for agents. Leaders should be thinking about the difference—and it starts with observability.

At Parable we continue to refine our structural approach to context graphs within our own platform—and this research only substantiates our conviction in their value for enterprise work.

A couple of key learnings worth sharing with technology leaders:

- **Dumping context into a prompt can actively hurt performance.**

Giving an agent a wall of organizational context before it starts working actually makes it worse at tasks it could have solved on its own (and be costly in tokens required to process). The quality and precision of context delivery matters enormously.

- **Agents that received no context outperformed agents that received bad context.**

When we built a threshold that suppressed low-confidence results and told the agent "I don't have a good match—explore on your own," it eliminated the worst failure modes. Silence is better than excess noise.

The implication for executives: more data is rarely the answer. What matters is the right data, delivered at the right moment, with clear confidence signals.

At Parable, we consistently advocate for returning to first principles—deeply understanding current workplace reality and patterns, redesigning workflows from that vantage point, then deploying agents. Taking the time to prepare and graph the data, implement baselines, and architect a measurement framework will yield better downstream returns.

What This Means for the Enterprise, and Where We're Headed

Enterprise observability—a clear, quantitative picture of operational reality that captures friction points, process debt, and redundancies—is a prerequisite for effective AI and agent deployment. It also serves as the foundation for a kind of context-aware, organizational GPS—a structured resource that agents can query and navigate.

Organizations are inherently messy and mutable, and context graphs should be not static artifacts constructed once at the outset of a transformation project. Our research points to an anticipated future state in which context graphs are dynamic entities, where the initial map is the starting point but evolves into a true system of understanding that learns and optimizes as it takes action within it, reinforcing relevant signals and letting stale context fade.

To extend the GPS metaphor further: a GPS can give an agent accurate directions; but real-world driving experience teaches the agent that a given road might flood during heavy rainstorms, or that a shortcut might work better at a certain time of day. Context-driven world models are that experience for agents—and building them requires that organizations invest deliberately in the infrastructure to make organizational data accessible, structured, and queryable.

The gap between agentic AI's promise and enterprise value creation does not stem from model shortcomings. We believe that it is fundamentally a context problem—but context is something organizations can actively build, structure, and leverage. Parable creates a map of how your organization works today, and over time, it becomes something much more: a dynamic system of understanding that learns, compounds, and makes every subsequent AI investment more precise. We've taken steps through this research to demonstrate just how much impact the right data and context can have on agentic knowledge work.