BURAK GALİBA
Thoughts on AI Transformation

ESSAY · ENTERPRISE AI

Everyone Has AI Agents. Almost Nobody Has Them in Production.

Most enterprises have adopted AI agents. Few run one in production — and Gartner expects 40% of agentic projects canceled by 2027. Mind the gap.

AGENTIC AI · PILOTS · GOVERNANCE

Burak Galiba · June 9, 2026 · 4 min read
https://burakgaliba.com/blog/the-agentic-ai-production-gap

SCAN TO READ ONLINE

Two numbers describe enterprise AI in mid-2026 better than any keynote.

Number one: in CrewAI's 2026 enterprise survey, 100% of enterprises — all of them — plan to expand agent adoption this year.

Number two: depending on whose data you trust, somewhere between 11% and 31% of enterprises have even *one* agent actually running in production. S&P Global and McKinsey put it around 31%; other surveys find 79% "adoption" collapsing to roughly 11% production. Industry observers have called the space between those numbers the largest deployment backlog in enterprise technology history.

Everyone is buying. Almost nobody is shipping. That gap is where the next two years of enterprise AI will be decided.

The stampede is real — and so is the cliff

The momentum numbers are genuinely staggering. Gartner expects 40% of enterprise applications to embed task-specific AI agents by the end of 2026, up from less than 5% in 2025. IDC forecasts agentic AI spending reaching $1.3 trillion by 2029 — more than 26% of all worldwide IT spending.

Now the cliff. The same Gartner that predicts the stampede also predicts that over 40% of agentic AI projects will be canceled by end-2027 — killed by escalating costs, unclear business value, or inadequate risk controls. Gartner also flags rampant "agent washing": chatbots and RPA scripts relabeled as agents to catch the budget wave.

Both predictions are about the same companies. The enterprises stampeding in without production discipline are the ones queued up for the 2027 cancellations.

Why agents get stuck between demo and deployment

An agent demo takes a week. An agent in production takes answers to questions the demo never asked:

- **Autonomy needs accountability.** A copilot suggests; an agent *acts*. The moment software executes decisions, someone must own every action it takes — legally, financially, operationally. Most organizations have no answer, so the agent stays in the sandbox.
- **Nobody built the evals.** What's the acceptable error rate for an agent touching customer refunds? Where's the threshold below which it escalates to a human? No eval framework, no security sign-off, no launch.
- **Risk controls are missing by design.** Pilots skip governance to move fast. But agents fail Gartner's cancellation test precisely on risk controls — the corner cut in month one becomes the cancellation memo in month eighteen.
- **Integration debt comes due.** Agents create value by acting across systems — which means authentication, permissions, and audit trails across every system they touch. That's real

engineering, and it was never in the pilot budget.

- **The value was never specified.** "Agentic" is a mechanism, not a business case. If nobody wrote down which cost line the agent moves, the CFO eventually asks — and that conversation is what a cancellation looks like.

How to be in the surviving 60%

The cancellation wave will not be random. It will select, with impressive precision, for projects that skipped the boring parts. The survivors will share a profile:

1. **One narrow, high-volume workflow** — not a general-purpose digital employee. Scope is the strongest risk control ever invented.
2. **A business metric with a baseline**, written down before development starts, owned by a named executive.
3. **Human-in-the-loop gates at every irreversible action** — payments, customer commitments, data changes — with autonomy expanded only as measured error rates earn it.
4. **Evals and audit trails from day one.** Log every action; test failure modes before launch, not after the incident.
5. **Governance sign-off as an entry gate, not an exit hurdle.** Teams that involve risk functions early report faster launches, not slower — the objections get engineered out instead of litigated at the end.

None of this is exotic. It's the same production discipline enterprises already apply to payment systems and ERP changes — applied to software that happens to reason.

The contrarian opportunity in the correction

Here's the part that should interest CEOs: the coming cancellation wave is a gift to disciplined operators.

While competitors burn 18 months and seven figures on agents that die in review, an enterprise that ships one governed, measured agent in a back-office workflow builds the muscle everything else will use. When the 2027 correction arrives, it will meet frozen budgets everywhere except the places with production numbers to defend.

The gap between adoption and production isn't a market failure. It's a ranking mechanism — and it's still early enough to choose your side of it.

Where to start

Pick your single most credible agent candidate and give it the five-item test above: narrow scope, baselined metric, human gates, evals, governance sign-off. Whatever it fails first is your actual project.

To see the whole board before you commit, my free AI Readiness Score takes roughly ten minutes — 20 questions spanning pilots, data, talent, and governance — and shows you whether your organization is built to land agents in production or to donate them to the 2027 cancellation statistics.

Read the essay online at https://burakgaliba.com/blog/the-agentic-ai-production-gap

I built a free AI Readiness Score — 20 questions across pilots, data, talent and governance, about ten minutes, no email required: https://burakgaliba.com/readiness