



MARTIN HURBAN
Manager, Technology Consulting
PwC



RICHARD DUDŽÁK
AI Engineer, Technology Consulting
PwC



WHY AGENTIC AI PROJECTS FAIL?

What we are observing in practice is a growing gap between the pace of innovation and the value realized. Agentic AI is advancing rapidly, yet many organizations are not seeing proportional business impact. PwC's May 2025 AI Agent Survey highlights strong momentum: 79% of executives report that they are already adopting AI agents, and 88% expect to further increase their AI investments.

The challenge is conversion. Teams often demonstrate technical feasibility through pilots and proofs of concept, but far fewer translate these into sustainable production outcomes. MIT's Project NANDA suggests only about 5% of custom enterprise AI solutions reach production. Gartner similarly forecasts that more than 40% of agentic AI initiatives will be discontinued by 2027 as costs outweigh perceived value.

In most cases, the limiting factor is not the technology itself, but the operating environment around it. Common constraints include unclear ownership, insufficient data readiness, lack of operational discipline, and weak user adoption. Based on our experience, these issues follow recurring patterns and are predictable, therefore largely avoidable.

In the following sections, we outline the five most

common patterns we encounter, along with a stage gated playbook designed to help organizations validate assumptions early, manage risk, and scale solutions only when there is clear evidence of value.

FIVE FAILURE PATTERNS

1. "Cool use case" with no business owner

Many AI projects begin with excitement and sponsorship, but without a single accountable owner for outcomes. Teams measure outputs ("we shipped a chatbot") rather than impact ("we reduced handling time by X"). For example: After delivering a document-drafting agent prototype, we learned the efficiency gains wouldn't justify the cost to bring it to production. This would have surfaced earlier if a value owner had been accountable for the KPI and clear "kill criteria" had been set upfront.

2. Data readiness is assumed, not proven

Teams often discover too late that critical data is missing, inaccessible at scale, or too inconsistent to use reliably. In many organizations, the "real" knowledge lives in PDFs, email threads, and unwritten know-how - costly to operationalize at scale. Example: In a use case processing hundreds of thousands of documents, an early data landscape assessment allowed a

realistic business case. Data access friction wasn't a hard blocker, but it significantly increased engineering effort and increased the cost curve.

3. Technology isn't mature enough for production deployment

Another trap is assuming capabilities the current stack can't deliver consistently. For example, fully autonomous agentic execution across systems or near-zero error tolerance might work in a pilot, but production reality forces scope changes. Example: In one use case, we achieved acceptable quality in a controlled pilot, but meeting production expectations for stability and consistency proved too costly. Targeted experiments exposed the gap early, avoiding unnecessary investment.

4. The solution isn't integrated into workflows

Even when outputs are good, adoption can stall if the tool lives outside the systems people use daily - or adds steps instead of removing them. Users treat it as "nice to have," and usage fades after the initial novelty. Example: For an internal support assistant, adoption rose when we embedded it in a channel employees already used daily rather than launching a separate app. Putting AI in the

Agentic AI projects fail when the business case is missing, the data foundation is unproven, and delivery and adoption aren't treated as a transformation. This article outlines five recurring failure patterns and a stage-gated playbook to scale successfully.

moment of work made the next step easier than the manual alternative, and usage (and value) followed.

5. Trust, risk, and compliance arrive too late

Rollouts stall when privacy, security, legal, or audit requirements aren't built in from the start. Or they "launch" but fail quietly because users don't trust the outputs enough to rely on them in daily work. Example: We've seen solutions stall late in delivery when control requirements surfaced after build, forcing rework and delays. We've also seen "launched" tools go unused when users didn't trust the outputs enough to rely on them. Building guardrails and clear escalation paths early makes compliance and trust part of the design.

A PRACTICAL PREVENTION PLAYBOOK: THREE GATES TO SCALE

AI initiatives fail most often at the conversion point, where teams can show feasibility in a pilot, but struggle to turn that into production impact. The gap is rarely effort, it's missing evidence at the moments where decisions should be made.

A stage-gated delivery path turns conversion into a managed process. It starts with exploration and prototypes, matures through MVP and pilot,

and only then commits to production and continuous improvement. Explore → Prototype → MVP → Pilot → Production → Operate & Improve shows how investment should increase in stages. Each phase is designed to answer a specific question, reducing uncertainty before you commit more time and budget.

Think of the gates as three moments to pause and make a deliberate decision: continue, adjust, or stop. The first gate tests clarity of intent, namely the business need, ownership, KPIs, and risk tolerance. The second tests feasibility, namely whether the end-to-end design, data, controls, integration, and performance expectations hold. The third tests economics and adoption, namely whether total cost of ownership is justified by measurable benefits and a realistic plan to drive usage.

SCALE EVIDENCE, NOT AMBITION

AI projects fail when organizations treat them as technology builds instead of outcome-driven change. The fix is simple, not easy: start with a measurable decision and an accountable owner, validate data and feasibility early, and design for compliance and trust from day one. Then scale in stages, only when evidence proves the next step is worth the investment.