Why AI Pilots Fail, and How Rapidly Growing and Mid-Sized Companies Should Start with Agentic AI

A practical guide to choosing, scoping, and delivering a first agentic AI project that reaches production

This whitepaper brings together key questions and answers from my executive briefings, conference speaking, participants in my training programs and additional research addressing process selection, AI automation, organizational readiness, and successful implementation.

James Proctor, The Inteq Group | November 2026

KEY TAKEAWAYS

1. Many AI pilots do not reach production or deliver sustained financial results. Common causes include unclear scope, undefined success criteria, poorly understood processes, and weak business ownership.
2. Task-level AI can improve productivity, but greater business value often requires redesigning workflows and defining which decisions agents should perform.
3. Improve the process before automating it. Clarifying rules, exceptions, and handoffs can reduce complexity and make AI easier to govern.
4. Choose the first process with three screens: routine judgment calls, steady volume, and outcomes you already measure.
5. Agree on decision boundaries and measurable success criteria before building. An accelerated first project can progress in weeks, but deployment timing depends on integration, testing, and risk.

How often do AI pilots fail?

Many AI pilots never reach production or deliver measurable business results. Research published in 2025 and 2026 shows that the challenge is not simply proving AI can work, but integrating it into real workflows and translating experimentation into sustained operational value.

- **Madrona** (August 2026)**:** In a survey of 150 senior enterprise decision-makers, 83% reported converting fewer than half of their AI pilots into production over the previous 12 months.
- **Dun & Bradstreet** (October 2026)**:** Across 10,000 surveyed businesses in 32 countries, 37% reported scaling AI pilots into production, while only 9% said most pilots had scaled with clear gains. Just 28% reported broad or strong AI ROI.
- **McKinsey & Company** (August 2026)**:** In a survey of 1,719 participants across 97 countries, 44% of organizations using AI reported scaling it across the enterprise, while 37% of all respondents reported an enterprise-level earnings impact from AI. High performers were much more likely to redesign workflows.

- **Gartner** (June 2025)**:** Gartner forecast that more than 40% of agentic AI projects would be canceled by the end of 2027, citing escalating costs, unclear business value, and inadequate risk controls. This is a forecast, not a measured cancellation rate.

These studies measure different outcomes: production deployment, scaling, financial returns, and projected cancellations. They should not be combined into a single failure rate. Although much of the evidence comes from large enterprises, rapidly growing and mid-sized companies face similar challenges, often with fewer resources to absorb unsuccessful initiatives.

Why do AI pilots fail?

Failed pilots tend to share the same handful of causes:

- **Scope that is too broad. Several processes, several departments, and a plan measured in quarters.**
- **Success that was never defined.** Without agreed measures and a baseline, there is no point at which the pilot is finished, so it either drifts or gets cut.
- **A process nobody fully understands.** The rules live in experienced people's heads, and the exceptions are handled from memory.
- **No business owner.** Nobody with authority can settle questions about rules, scope, and limits, so decisions wait.
- **Chosen for visibility, not fit.** The pilot demonstrates something impressive but doesn't connect to an outcome leadership tracks, so there is no case for funding the next step.

Gartner's analysts add a sixth: many use cases labeled agentic don't need an agent at all. Simpler automation would do the job at lower cost and risk. Selection discipline catches that early.

Why doesn't adding AI to tasks move the business?

Copilots and chat assistants help individuals draft, summarize, and search faster. That's useful, but the gains level off, because the process around the person hasn't changed. An agent is different. It becomes part of the process: it reads the incoming request, applies your rules, makes the routine decision, updates the system, and sends exceptions to the right person.

The same logic explains why task automation stalls. A task is one step, such as extracting data from an invoice. A process is the full flow of work, including the decisions, exceptions, and handoffs between people. Most of the delay in a process isn't in the steps. It's between them: waiting for an approval, chasing missing information, sending an exception back and forth. Speeding up one step doesn't clear a queue that forms at the next handoff.

McKinsey's 2026 State of AI survey reinforces this distinction. Nearly three-quarters of AI high performers said they had fundamentally redesigned workflows, compared with approximately one-quarter of other respondents. This is an association, not proof that redesign alone causes better results. The model is one component; changing how work flows is central to capturing value.

Should you reengineer a process before automating it?

Automating the process you have today means the agent inherits every flaw in it. Most processes have accumulated workarounds, redundant steps, unclear rules, and exceptions that experienced people handle by judgment. People absorb those problems. An agent executes them, faster, at volume, and with confidence.

Improving the process before automation can also reduce implementation complexity and operating costs. Removing unnecessary steps, clarifying decision rules, and eliminating redundant handoffs may reduce the work an agent must perform, potentially allowing simpler automation, smaller models, and fewer decisions requiring oversight. This is business process reengineering, applied to one process at a time, with emphasis on who makes each decision, using what information, and within what authority.

> ***You can iterate your way to a better screen. You can't iterate your way to an authority boundary.***

For leaders trained on Agile delivery, this can sound like big design up front. It isn't. Iteration helps people recognize inconsistencies, apply experience, and adjust their approach. AI agents can detect some anomalies, but they cannot be assumed to recognize every business exception or independently determine the limits of their authority. Defining what an agent may decide, within what limits, and where exceptions go is a focused preparation step that should precede deployment.

How do you choose the first agentic AI process?

Choose a process where skilled people spend a lot of time on routine judgment calls, the volume is steady, and the outcome is already measured. Those three screens remove most poor candidates quickly:

- **Routine judgment calls.** Reviewing, classifying, approving, or routing, where rules can be stated and the exceptions are known.
- **Steady volume.** Dozens or hundreds of cases a day, enough that improvement shows up in the numbers.
- **Outcomes you already measure.** Cycle time, cost per transaction, error rates, or backlog, with a baseline you can record before work begins.

These three screens identify promising candidates. Before selecting the first project, also assess data readiness, integration complexity, decision risk, and access to business expertise. A high-value process may still be a poor first candidate if its implementation risks are disproportionate to its expected benefits.

Common candidates in mid-sized companies include order entry and order exceptions, quoting and estimating, customer service triage, accounts payable matching, onboarding, and collections. The list is a starting point, not a prescription. Screen candidates against your own process profile rather than picking from an industry list.

Also screen out the red flags: processes nobody can explain clearly, processes where every case is unique, politically sensitive processes, and anything chosen mainly for visibility. Check the data the process depends on, not the whole company's data. If the agent will check orders against pricing, the pricing data must be accurate. Customer records that the process doesn't rely on can wait.

A worked example: a manufacturer screening three candidates might find that its annual strategic plan involves judgment but happens once a year, and that supplier data entry has volume but needs only simple automation. Order exceptions pass all three screens. That is the first project.

What should be agreed before an AI agent is built?

Two things should be signed off by the business owner before building begins. The first is the limits: which decisions the agent will handle, what it is not allowed to do, and where uncertain cases go. The second is the finish line: the share of cases it should handle without a person, the accuracy required, and the response time. Together, they are the definition of done for the agent. They are what stops a pilot from running indefinitely without a decision.

Setting the limits depends on capturing how your experts actually decide. Generic interviews rarely work, because experienced people struggle to describe judgment in the abstract. Walking through real cases does. For each decision, record what the person looked at, what they concluded, why, and which cases they would treat differently. The exceptions are where the knowledge is. The result is a set of explicit decision rules that an agent can execute, a new hire can learn, and an auditor can follow.

How long does a first agentic AI project take?

For a well-scoped first process, an initial capability can sometimes be developed and tested in weeks rather than quarters. Timing depends on integration complexity, data readiness, security requirements, testing, and the decisions the agent will be authorized to make. The following four-week sequence illustrates an accelerated approach, not a guaranteed production schedule:

- **Week 1: Confirm and baseline.** Confirm the process, name the owner, and record the current performance.
- **Week 2: Map the decisions.** Sit with the people who do the work, including the unwritten rules.
- **Week 3: Set limits and criteria.** Agree on what the agent may do and what success means.
- **Week 4: Validate and launch cautiously, if ready.** Test on representative historical and exception cases, complete required reviews, and introduce monitored use with human oversight.

An agent should begin conservatively, handling cases that meet defined decision criteria and escalating unusual situations to a person. Its authority should expand only after performance is validated on representative cases, including exceptions and higher-risk situations, with ongoing monitoring. If an initial project is expected to take several quarters, review the scope, dependencies, and implementation requirements rather than assuming the scope is necessarily too broad.

How does a first project become an AI roadmap?

A successful first project produces a reusable method, not just one automated process: decision mapping, controls, a testing approach, and technical connections that make the second process faster than the first.

Build the roadmap from your processes, not from technology trends. Inventory the candidates, score each on value, readiness, and risk, and sequence them so each project reuses what the last one built. Each should have an owner, a baseline, and success measures. A sequence of projects measured in weeks is easier to defend, and to adjust, than a multi-year program.

Why are rapidly growing and mid-sized companies well placed to get this right?

Rapidly growing and mid-sized companies can have meaningful advantages when implementing agentic AI. Shorter decision chains, closer access to business expertise, and organizational agility may make it easier to establish ownership and move from analysis to implementation. These advantages depend on the company's process maturity, systems, and governance. The strongest starting point remains a deliberately selected process with clear ownership, measurable outcomes, and manageable implementation risk.

Questions answered in this paper

Each question below has a short, direct answer in Getting Started with Agentic AI: Top 10 Questions Answered for Rapidly Growing and Mid-Sized Businesses.

- Why do so many AI pilots fail to make it into production?
- We already use ChatGPT, Claude, etc. How is agentic AI different from this?
- What's the difference between automating a task and automating a process?
- Why not just automate the process we have today?
- Is business process reengineering still relevant in the age of AI?
- Where should we start? Which process?
- What are the best AI agent use cases for a mid-sized company?
- How clean does our data need to be before we use AI agents?
- How do you turn an expert's judgment into rules an AI agent can follow?
- How do we build an AI roadmap our leadership team can defend?

Sources

1. Madrona, "Harnessing Enterprise Value – The ROI of AI" (August 13, 2026). https://www.madrona.com/harnessing-enterprise-value-roi-of-ai/
2. Dun & Bradstreet, "AI Momentum Survey" (October 28, 2026). https://www.dnb.com/en-us/newsroom/press-releases/dnb-survey-finds-ai-returns-stable-for-most.html
3. McKinsey & Company, "The State of AI in 2026: On the Road to ROI" (August 25, 2026). https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai
4. Gartner, "Over 40% of Agentic AI Projects Will Be Canceled by End of 2027" (June 25, 2025). https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027
5. RAND Corporation, "The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed" (August 2024); discusses estimates from prior research. https://www.rand.org/pubs/research_reports/RRA2680-1.html

About The Inteq Group

The Inteq Group helps organizations strengthen business performance through business process reengineering, business systems analysis, and agentic AI solutions built on open-weight technology deployed in infrastructure the client controls.

inteqgroup.com | 800-719-4627 | info@inteqgroup.com