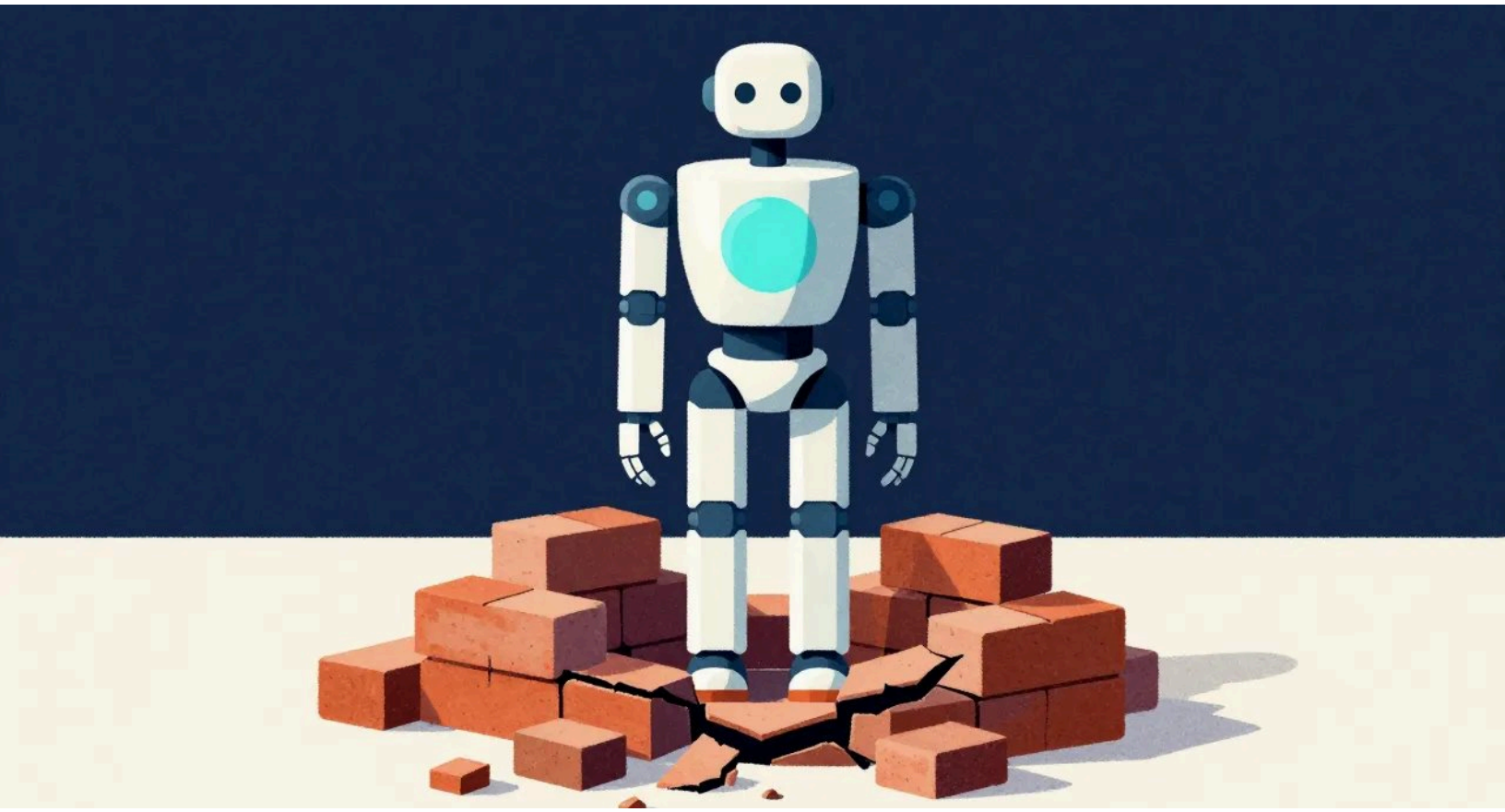


EXECUTIVE PAPER · AI

AI Projects Fail on Their Data

Most stalled AI programs were funded in the wrong order. Executives who pay for discovery and data foundations before agents get projects that reach production.

AUTHOR

Chris Watkins

SERVICE

AI

FOR

CEO, CFO · 12 min read

AT A GLANCE

The problem. Enterprise AI spending has outrun results. MIT researchers reported in 2025 that 95% of organizations saw no measurable business return from generative AI, and S&P Global found 42% of companies had abandoned most of their AI initiatives that year, up from 17% the year before. The models performed in the demo. The projects stalled when they met real company data and real production systems.

Why it persists. AI budgets are approved one use case at a time, so each pilot pays to find, clean, and connect its own data and leaves little behind for the next one. The model is the visible line item. Data and platform work stays invisible until it blocks a launch.

The recommendation. Fund AI in sequence. Start with a short discovery phase that picks problems worth solving. Build a shared data foundation with named owners and quality targets, then infrastructure that can be rebuilt and audited, then agents. Release money for each stage only when the stage before it passes a written exit test.

The Decision in Front of You

Most leadership teams now carry an AI budget line and a queue of agent proposals. The decision here is the order in which that money is spent.

The common order puts agents first. A business unit picks a use case, a vendor or internal team builds a pilot, and data problems get handled as they surface. The alternative spends the first months, and a real share of the budget, on deciding which problems matter and making the underlying data trustworthy. Agents come after.

Gartner estimated in 2024 that deploying generative AI costs an enterprise between $5 million and $20 million depending on the approach, and predicted that 30% of generative AI projects would be abandoned after proof of concept by the end of 2025 [5]. Poor data quality led its list of reasons [5]. In June 2025 Gartner forecast that more than 40% of agentic AI projects will be canceled by the end of 2027 because of rising costs, unclear business value, or inadequate risk controls [6].

What's Happening

Adoption is broad and returns are thin

McKinsey's 2025 global survey of 1,993 respondents found 88% of organizations using AI in at least one business function [11]. About a third had begun to scale it. Only 39% could attribute any EBIT impact to AI, and most of those put the effect below 5% [11]. Roughly 6% of respondents qualified as high performers.

MIT's Project NANDA published "The GenAI Divide: State of AI in Business 2025" after reviewing more than 300 public AI initiatives and conducting 52 interviews and 153 surveys between January and June 2025 [1]. The authors estimated enterprise spending on generative AI at $30 billion to $40 billion, with 95% of organizations reporting no business return [1]. Among organizations that evaluated enterprise-grade AI systems, 60% reached evaluation, 20% reached a pilot, and 5% reached production [1]. The Register's coverage of the same report noted that only 5% of custom enterprise AI tools reach production [2]. The sample is small and drawn largely from conference attendees, so treat the 95% as a direction of travel.

In S&P Global Market Intelligence's 2025 survey of more than 1,000 respondents in North America and Europe, 42% of companies had scrapped most of their AI initiatives, against 17% the year before [7]. The average organization abandoned 46% of its proofs of concept before they reached production [7].

The failures trace to data and plumbing

The research agrees on where projects break. RAND interviewed 65 experienced data scientists and engineers in 2023 and published five root causes of AI project failure in 2024 [3]. Two concern foundations directly: organizations lacking the data needed to train an effective model, and lacking the infrastructure to manage data and deploy finished models [3]. A third, misunderstanding the problem the AI is supposed to solve, is the failure that discovery work exists to prevent. RAND cites estimates that more than 80% of AI projects fail, roughly twice the rate of IT projects without AI [3].

MIT's authors found the divide had little to do with model intelligence or raw infrastructure capacity [1]. The core barrier they named was systems that do not retain feedback, adapt to context, or fit day-to-day workflows [1]. Inside a company, that context is data: customer history, contract terms, pricing rules.

Data leaders say the same about their own organizations. In a July 2024 Gartner survey, 63% of organizations lacked confidence in their data management practices for AI, and Gartner predicted that 60% of AI projects without AI-ready data will be abandoned through 2026 [4]. IBM's 2025 survey of 1,700 chief data officers found only 26% were confident their data could support new AI-enabled revenue streams [8]. Informatica's 2026 survey of 600 data leaders found 57% named data reliability as a key barrier to moving AI from pilot to production [10]. Informatica sells data management software.

Governance shows up in the margin

IBM's 2026 study of 2,000 technology executives found 77% saying AI adoption was outpacing their governance [9]. Organizations that built governance and financial controls into their AI systems reported 25% fewer security incidents than those relying on manual approval [9]. They also deployed 16 times as many agents and reported operating margins 18% higher [9]. These are self-reported correlations, and they point the same way as the failure research.

Two recent cases

New York City's business chatbot, 2023 to 2026. The city launched an AI chatbot in fall 2023 to answer business owners' questions about city rules [12]. Testing by The Markup and THE CITY in March 2024 showed it telling users that landlords could refuse Section 8 vouchers, that employers could take a cut of workers' tips, and that businesses could refuse cash, all contrary to city law [12]. The city added disclaimers and narrowed what the bot would answer. In January 2026 the new mayor called it "functionally unusable," and the city took it down in early February [12]. Building its foundations reportedly cost close to $600,000, and maintenance costs were never disclosed [12]. Reporting does not say what data the bot drew on. Two years of patching a deployed system never made its answers reliable enough to keep.

Deloitte and the Australian government, 2025. Australia's Department of Employment and Workplace Relations paid Deloitte A$440,000 for a review of a welfare compliance system [13]. After the report was published in July 2025, a University of Sydney academic found more than a dozen nonexistent references and a fabricated quote attributed to a federal court judgment [13]. The corrected report disclosed that a GPT-4o model had been used to fill traceability and documentation gaps, and Deloitte agreed to refund the final installment of the contract [13]. The model was asked to supply what the source record lacked, and no step checked its output against that record.

Why the Usual Approach Fails

The usual approach budgets AI the way companies budget software features: one use case, one sponsor, one project.

Each pilot pays for its own data work and keeps it private. The team building an invoice agent extracts vendor records, reconciles duplicate suppliers, and negotiates access with three system owners. Six months later the team building a procurement agent does the same work again on the same records. Neither project's budget shows that money as data spending, so nobody sees the total.

Pilots run on curated data. A proof of concept usually runs on an extract someone cleaned for the demo. Production means live systems, missing fields, conflicting definitions of "customer" across departments, and access rules nobody wrote down. Gartner's analyst noted that hype obscures the real cost and complexity of deploying agents at scale [6].

Infrastructure cannot be rebuilt. Many pilots live in a notebook or one engineer's cloud account. When the model version changes or the engineer leaves, nobody can reproduce the result or prove which data produced a given answer. RAND names inadequate infrastructure for managing data and deploying models as a root cause [3].

Data has no owner. The sponsor owns the agent. Nobody owns the customer master, the price file, or the contract repository the agent reads. When the agent gets something wrong, the investigation stops at "bad data" because no one is accountable for fixing it. McKinsey's high performers were nearly three

times as likely as others to have fundamentally redesigned workflows [11], and that kind of redesign needs process and data owners involved from the start.

The Framework: Four Funding Gates

Treat AI as four stages, each with a written exit test, and release money for the next stage only when the current one passes.

Exhibit 1. Four funding gates: agents get money last, and every later agent reuses the first three stages



Source: author's framework, informed by RAND's root causes of AI project failure [3] and Gartner's guidance on AI-ready data [4].

Gate 1: Discovery. Spend four to eight weeks listing candidate use cases with their dollar value, an accountable process owner, and the systems each would read. Exit test: three ranked use cases, each with a dollar value the CFO accepts, a named owner, and a list of the data it needs.

Gate 2: Data foundation. For the data those use cases share, assign an owner to each dataset, agree on definitions, set measurable quality targets, and write down who may see what. Limit scope to the fields the first use cases need. Exit test: every dataset in scope has an owner, and measured quality meets the agreed target. The companion paper in this series, "Define Your Data Before You Deploy an Agent," covers what this layer contains.

Gate 3: Platform. Build infrastructure that can be recreated from source: environments defined in code, versioned data snapshots, versioned prompts and model settings, an evaluation suite, cost tracking, and a separate identity for each agent. Exit test: the team can rebuild any environment in a day and trace any AI output to the data and model version that produced it.

Gate 4: Agents. Build the first agent on the shared base. Exit test: accuracy measured against a test set the process owner signed off, and a known production cost per transaction. Each later use case starts at Gate 4 and adds only the data it needs that the foundation does not yet hold.

Where the money goes

The following is an illustrative model for a hypothetical mid-size distributor planning five AI use cases over two years. Substitute your own figures.

Exhibit 2. Illustrative two-year cost: foundation-first breaks even at the third use case

Assumption or result	Agent-first	Foundation-first
Shared data and platform foundation (one time)	$0	$600,000
Data preparation per use case	$300,000	$80,000
Agent build per use case	$200,000	$200,000
Cumulative cost after 1 use case	$500,000	$880,000
Cumulative cost after 3 use cases	$1,500,000	$1,440,000
Cumulative cost after 5 use cases	$2,500,000	$2,000,000
Difference after 5 use cases		$500,000 lower

Source: author's illustrative estimate. Assumes each agent-first project repeats extraction, cleanup, access review, and integration for overlapping data at $300,000 per project; a shared foundation costs $600,000 once and cuts incremental data work to $80,000 per use case; agent build cost is equal in both paths. The model excludes the cost of failed projects, which the failure research above suggests is higher on the agent-first path.

The model assumes every agent-first project succeeds. Each one that stalls in pilot widens the gap by the money spent on it.

Costs, Risks, and Tradeoffs

Exhibit 3. Tradeoffs of funding foundations before agents

Area	Cost or risk	Mitigation
Time to first agent	The first production agent arrives a quarter or two later than on the agent-first path	Run discovery and a narrow data foundation in parallel; ship one internal, low-risk assistant early to keep momentum
Upfront spend	The foundation costs more than the first pilot would have, as Exhibit 2 shows	Scope Gate 2 to the fields the first three use cases need, and report reuse as each new use case starts
Business unit patience	Sponsors see money going to data work they cannot demo	Publish the ranked use-case list and dates; show quality metrics improving monthly
Scope creep	Data programs can expand into multi-year enterprise cleanups	Tie every data task to a named use case; anything without one waits
When this is wrong	A company with one narrow use case, or one whose data is already governed with owners and quality measures, gains little from a separate foundation phase	Go straight to Gate 4 with a short discovery check

Source: author's analysis.

How to Act

Exhibit 4. A one-year roadmap

Window	Actions	Owner	Milestone
First 30 days	Inventory every AI pilot, its spend to date, and the data it reads. Require new AI proposals to pass Gate 1 before funding.	CFO with CIO or CTO	Complete list of AI spend; gating rule in the approval process
First 30 days	Name one executive accountable for the AI program and a process owner for each candidate use case.	CEO	Accountable executive named; owners confirmed in writing
First 90 days	Complete discovery. Rank use cases by dollar value and data overlap.	AI program lead with business unit heads	Three ranked use cases with values the CFO signs off
First 90 days	Assign owners to the shared datasets. Measure quality on the fields in scope. Document access rules.	Chief data officer or data lead	Baseline quality report; every in-scope dataset has an owner
First 90 days	Stand up a reproducible platform: environments in code, versioned data and prompts, evaluation suite, per-agent identity.	Platform or infrastructure lead	Any environment rebuilt from source in under a day
First 365 days	Ship the first two agents on the shared foundation. Track reuse for each new use case. Report AI spend split between reusable foundation and single-use work.	AI program lead with CFO	Two agents in production with signed-off accuracy; foundation reused by every new use case; quarterly spend report to the board

Source: author's recommendation.

Questions to Ask Your Team

1. How much have we spent on AI pilots in the last 18 months, and how many are in production today?
2. For each pilot, what share of the effort went into finding, cleaning, and connecting data?
3. Which datasets do two or more of our AI proposals depend on, and who owns each one?
4. If a regulator asked how an AI output was produced last March, could we rebuild the exact data and model that produced it?
5. What is the measured error rate on the data our first agent will read?
6. What written test does a project pass before it moves from pilot to production funding?

Closing

The research on stalled AI projects points to data and infrastructure far more often than to the model. Companies that fund agents first pay for data work repeatedly, inside budgets that hide it, and lose much of it when pilots die. Fund discovery first, then a shared data foundation with owners and quality targets, then infrastructure that can be rebuilt and audited, and give agents money last. Each agent after the first will cost less and reach production sooner.

Sources

[1] Virtualization Review. "MIT Report Finds Most AI Business Investments Fail Reveals GenAI Divide." Coverage of MIT Project NANDA, "The GenAI Divide: State of AI in Business 2025." 2025. https://virtualizationreview.com/Articles/2025/08/19/MIT-Report-Finds-Most-AI-Business-Investments-Fail-Reveals-GenAI-Divide.aspx.

[2] The Register. "GenAI FOMO has spurred businesses to light nearly $40 billion on fire." Coverage of MIT Project NANDA report. 2025. https://www.theregister.com/2025/08/18/generative_ai_zero_return_95_percent.

[3] Ryseff, J., De Bruhl, B. F., and Newberry, S. J. "The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed." RAND Corporation, RR-A2680-1. 2024. https://www.rand.org/pubs/research_reports/RRA2680-1.html.

[4] Freevacy. "Gartner reveals 60% of AI projects with data issues will fail." Coverage of Gartner press release, "Lack of AI-Ready Data Puts AI Projects at Risk." February 26, 2025. https://www.freevacy.com/news/gartner/gartner-reveals-60-of-ai-projects-with-data-issues-will-fail/6155.

[5] Intelligent CIO. "Gartner predicts 30% of generative AI projects will be abandoned after proof of concept by end of 2025." Coverage of Gartner press release. August 5, 2024. https://www.intelligentcio.com/eu/2024/08/05/gartner-predicts-30-of-generative-ai-projects-will-be-abandoned-after-proof-of-concept-by-end-of-2025/.

[6] Outlook Business. "Over 40% of Agentic AI Projects Will Be Scrapped by 2027, Says Gartner." Coverage of Gartner press release, June 25, 2025. 2025. https://www.outlookbusiness.com/artificial-intelligence/over-40-of-agentic-ai-projects-will-be-scrapped-by-2027-says-gartner.

[7] CIO Dive. Coverage of S&P Global Market Intelligence survey on AI project abandonment. March 14, 2025. https://www.ciodive.com/news/AI-project-fail-data-SPGlobal/742590/.

[8] IBM. "IBM Study: Chief Data Officers Redefine Strategies as AI Ambitions Outpace Readiness." Press release, IBM Institute for Business Value CDO Study. November 13, 2025. https://newsroom.ibm.com/2025-11-13-ibm-study-chief-data-officers-redefine-strategies-as-ai-ambitions-outpace-readiness.

[9] CIO Dive. Coverage of IBM Institute for Business Value, "2026 Tech Leader Study." June 2026. https://ciodive.com/news/tech-leaders-ai-deployment-underprepared/822295.

[10] Informatica. "New Global CDO Report Reveals Data Governance and AI Literacy as Key Accelerators in AI Adoption." Press release via Business Wire, CDO Insights 2026. January 27, 2026. https://www.businesswire.com/news/home/20260127455879/en/New-Global-CDO-Report-Reveals-Data-Governance-and-AI-Literacy-as-Key-Accelerators-in-AI-Adoption.

[11] McKinsey & Company. "The state of AI in 2025: Agents, innovation, and transformation." 2025. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2025.

[12] The Markup. "Mamdani to Kill the NYC AI Chatbot We Caught Telling Businesses to Break the Law." January 30, 2026. https://themarkup.org/artificial-intelligence/2026/01/30/mamdani-to-kill-the-nyc-ai-chatbot-we-caught-telling-businesses-to-break-the-law.

[13] The Register. "Deloitte refunds Aussie gov after AI fabrications slip into $440K welfare report." October 6, 2025. https://www.theregister.com/2025/10/06/deloitte_ai_report_australia/.

The views in this paper are the author's own and do not represent any employer. Scenarios are hypothetical. This paper is not legal, financial, or compliance advice.