WHITE PAPER

Before You Fund Another AI Pilot

Five leadership beliefs that make AI activity look like progress, and the five conditions that turn investment into value

A leadership decision guide for mid-market executives working to turn AI activity into measurable business value.

JR Key Advisory

JR Key · JR Key Advisory
Portland, Oregon

Activity is easy to create

AI activity is easy to create. A leadership team can approve tools, encourage experimentation, launch pilots and ask every department to find use cases, and still lack a clear answer to the question that matters:

What is becoming more effective because of this investment?

The pressure is real. Competitors are experimenting, employees are asking for tools, and stakeholders expect a point of view.

For a $10M–$100M company, that question arrives quickly. A company this size can often move faster than an enterprise, but it has less room for scattered pilots, unclear ownership or weak evidence. Each pilot also costs more than its budget. It takes leadership attention and the patience of the people asked to change how they work, and every pilot that fades without a clear result makes the next AI proposal harder to approve.

Pilots look like a reasonable place to start because they seem small. But every pilot is designed around assumptions about what AI is for and where its value will show up. When those assumptions go unexamined, a pilot can succeed on its own terms and still not move the business.

This is a familiar pattern. Business process reengineering, ERP and digital transformation each carried a version of the same promise: new technology, new processes, better performance. Each wave was different, but the same management lesson kept resurfacing. Adoption of a system isn't the same as validation of the strategy behind it.

Adoption of a system isn't the same as validation of the strategy behind it.

AI is different in speed and reach. Employees can experiment without waiting for a formal project, vendors add AI features faster than leadership teams can evaluate them, and work that once needed a specialist can now be roughed out by almost anyone in minutes.

The visible results haven't kept pace with the visible activity. Deloitte found that only a quarter of the leaders it surveyed had moved 40% or more of their AI pilots into production.[1] KPMG found established returns far more often where AI costs were visible and someone was accountable for outcomes.[2] BCG and Gartner each found that when no one decides where time saved by AI should go, it tends to be absorbed by lower-value work.[3],[4]

A WORD ON THE EVIDENCE

Public, measured case studies of AI at $10M–$100M companies are scarce. Most of the named cases here come from large companies. Read them as warning signals: each shows a leadership pattern worth recognizing before your next AI move. Each section closes with a translation for a mid-market business.

AI activity is not AI progress. Before the next pilot receives budget, it's worth examining the beliefs behind the activity you already have.

The five conditions for AI progress

Five conditions shape whether AI investment becomes measurable progress.

01 **Direction.** What should AI help the business become better at, before tools, pilots or partners define the path?

02 **Decisions.** Which AI opportunities deserve budget, capacity and executive attention before more experiments are added?

03 **Work.** Which workflows, handoffs, decisions and review points need to change before AI can improve how work gets done?

04 **People.** How should AI increase the capacity and capability of the people already in the business, before workforce reduction becomes the default value story?

05 **Progress.** What evidence will show that AI improved the business, before activity metrics are treated as value?

Each condition pairs with a belief that sounds reasonable but leaves something out. When leadership acts on an incomplete belief, AI activity tends to grow faster than the organization's ability to judge whether it's working.

Together, these conditions provide a practical screen for leadership teams to make the next AI investment easier to justify, easier to govern, and easier to evaluate.

01

Direction

What should AI help the business become better at, before tools, pilots or partners define the path?

THE BELIEF

"Choosing the right AI tool or platform is the strategy."

Tool selection is concrete. It has a price, a contract and a launch date, which makes it easy to fund and report, and it answers a real pressure. A Gartner session in September 2026 described CIOs as "under heavy internal and external pressure to increase AI adoption,"[5] and a signed platform agreement looks like a response to that pressure.

The belief breaks when the available tool starts to decide what AI is for. A platform shapes capability, cost, integration and adoption. It can't tell you which customer, operational, or growth priority comes first. When leadership hasn't made that choice, the platform's features tend to make it instead.

When leadership hasn't decided what AI is for, a platform's features tend to decide it instead.

In many companies, that's already happened. By the time AI reaches the leadership agenda, tools are in use, someone has run a pilot, and teams have bought their own subscriptions. Those purchases are now the closest thing the organization has to an AI investment decision.

RAND's interviews with 65 experienced data scientists and engineers found that the most common root cause of failed AI projects was stakeholders misunderstanding or miscommunicating the problem the system was meant to solve.[6] A tool can't clarify that problem for you. In a randomized field experiment with 515 firms, Kim, Kim and Koning found that companies prompted to look for AI use cases across a wider set of functions found 44% more use cases, completed 12% more tasks and were 18% more likely to acquire paying customers.[7] It's a working paper based on high-growth startups, but it's a causal result, and it suggests looking across the business before a single tool narrows the conversation.

Important nuance. Buying is often the right move. MIT researchers reported in 2025 that purchased tools and partnerships succeeded about twice as often as internal builds.[8] Buying can reduce complexity and speed deployment. That strengthens the case for setting direction first, because faster deployment only pays off once leadership has agreed what AI is for.

MID-MARKET TRANSLATION

The evidence	Adoption pressure is rising. Firms prompted to look for AI uses across more of the business found more of them, and failed AI projects most often trace back to a misunderstood problem.
What it reveals	Tool decisions become a stand-in for direction when the outcome isn't agreed on first.
At $10M–$100M	The team can end up shaping its AI plans around what a tool can do, before agreeing on what the business needs it to do.
The leadership question	**What business direction are we asking AI to support before we evaluate the next tool or platform?**

02

Decisions

Which AI opportunities deserve budget, capacity and executive attention before more experiments are added?

THE BELIEF

"More AI pilots means more progress."

Pilots feel prudent. They're small, easy to approve, and they keep the company learning. That logic only holds up when the portfolio has rules for deciding what happens next.

The belief breaks when leaders can't tell learning activity from investment progress. A pilot answers a question only if someone set the question in advance. Before pilots multiply, leadership needs agreement on what counts as success, who owns the result, which process or measure should change, what evidence would justify scaling, stopping or redesigning, and how much cost, capacity and attention the pilot is allowed to consume.

Gartner's September 2026 survey of 1,303 respondents at organizations with at least $50 million in revenue found that only 22% had scaled AI across multiple business units. The organizations doing better "constantly track ROI of AI initiatives, treat AI as a portfolio of value," and "regularly assess project performance and reallocate or discontinue underperforming initiatives." Roughly 11% didn't know what their function had spent on AI in 2025.[9]

Deloitte's 2026 survey of 3,235 leaders points the same way: "only 25% of respondents have moved 40% or more of their AI pilots into production," and 37% reported "only using AI at a surface level with little or no change to underlying business processes."[1] KPMG's Q2 2026 pulse found that 49% "have delayed or scaled back AI agent deployments when costs started outweighing benefits," and that established ROI showed up at 15% of organizations with full visibility into AI costs against 3% without, and at 14% with clear accountability for AI outcomes against 4% without.[2]

"We learned early that twenty pilots do not equal one transformation."

Marianne Johnson, Cox Automotive, as quoted in a Fortune op-ed[10]

No source shows that pilot count itself drives or blocks value. Established returns simply show up more often where costs are visible and someone owns the outcome.

Stopping is the discipline that makes the rest work. A company that can't stop an AI initiative can't really prioritize one. A bounded pilot with a stopping date is a test, a small commitment that can be ended if it doesn't prove its value. A simple check is to name an AI initiative the company deliberately stopped, and the criteria that stopped it. If nothing comes to mind, initiatives are being launched, but the organization isn't yet choosing among them.

A company that can't stop an AI initiative can't really prioritize one.

In regulated work, this sequence often isn't optional. If an AI system touches a decision that a regulator, auditor or court could later ask you to justify (like credit, hiring, clinical support) its purpose, ownership, boundaries and validation need to be settled before deployment. Legal and compliance teams can test a system against those decisions, but they can't supply the business purpose. If leadership hasn't decided it, whoever built or bought the system has.

Important nuance. Pilots are useful when they're designed to answer a real decision. They reduce risk and expose constraints. The trouble starts when pilots stand in for decision quality. Before adding another pilot, it helps to agree on what would make the team proceed, what would make it pause, and what evidence would justify further investment.

MID-MARKET TRANSLATION

The evidence	Few pilots reach production, and reported returns are higher where cost visibility and accountability exist.
What it reveals	Pilot volume is a weak signal without a way to decide what continues.
At $10M–$100M	The same few leaders and subject-matter experts run the business and staff the pilots. Loosely scoped pilots pull them away from that work and into projects that never become operational.
The leadership question	**What criteria determine which AI opportunities deserve budget, capacity, and executive attention before another pilot is approved?**

03

Work

Which workflows, handoffs, decisions and review points need to change before AI can improve how work gets done?

THE BELIEF

"Adding AI to existing workflows will produce the gains."

Adding AI to the current workflow is easier than redesigning it. It produces practical-feeling use cases and postpones harder conversations about roles, handoffs, approvals and risk.

The belief breaks because few existing workflows were designed with AI in mind. When AI is added, the same handoffs, approvals and review steps stay in place, so one step speeds up while the real constraint stays put. A team drafts proposals faster and still waits on the same legal review. A service group summarizes cases faster and still escalates every exception to the same senior specialist. The leadership work is deciding where AI should change the work and where human judgment should stay central.

McKinsey's 2026 survey of 1,719 organizations found that "nearly three-quarters of high performers report fundamentally redesigning workflows," against "just one-quarter of other respondents."[11] High performers are defined by their reported outcomes, so this is an association, not proof that redesign caused the results. PwC's April 2026 study of 1,217 senior executives found that 74% of AI's economic value is captured by 20% of organizations, and that those leaders were "twice as likely to redesign workflows to incorporate AI rather than simply adding AI tools."[12]

Allianz shows what the alternative looks like. Instead of speeding up steps in its existing claims process, it rebuilt food-spoilage claims in Australia around seven task agents and kept payout decisions with people by design. The company reports an 80% reduction in processing and settlement time.[13] Swiss Re Corporate Solutions' Tina Baacke describes the same approach as "redesigning workflows end-to-end rather than automating isolated tasks."[14]

Important nuance. Redesign doesn't have to start big. A small, low-stakes workflow rebuilt end to end with AI in mind is often a better place to start learning than adding AI to existing processes one step at a time, because it shows which handoffs, decisions and review points need to change. Existing workflows still matter as a map of where work, risk and customer impact actually live.

MID-MARKET TRANSLATION

The evidence	The organizations capturing the most value from AI are far more likely to redesign workflows than to add tools to existing ones.
What it reveals	Value is shaped by the design of the surrounding work, including where judgment, review, automation and accountability belong.
At $10M–$100M	Experienced people often bridge process gaps by hand. If AI is added before those gaps are surfaced, work moves faster but the rework, delays and risk stay.
The leadership question	**Which workflows, decisions, handoffs, and review points need to change before AI can improve the business?**

04

People

How should AI increase the capacity and capability of the people already in the business, before workforce reduction becomes the default value story?

THE BELIEF

"The main value is doing the same work with fewer people."

Labor efficiency is the easiest AI value story to explain, and leadership teams understand it quickly. Before it becomes the dominant story, leadership should know what's at risk and what higher-value capacity AI could create.

The belief breaks when headcount becomes the main lens before the work, the quality standard, the customer expectation, and the knowledge base are understood. Capacity returned by AI could go to better decisions, faster response, higher quality, deeper customer work, or growth work that's been under-resourced. In mid-market companies, a lot of the institutional knowledge lives in people's heads and nowhere else. Experienced people know which customer exception matters, where the process bends, and which handoff fails under pressure. Reducing staff before understanding that can carry costs that only appear later.

Expectations seem to have run ahead of results. In McKinsey's 2026 survey, 14% of AI-using organizations said AI had contributed to a decline in workforce size in the past year, compared with 32% in its 2025 survey who expected workforce reductions.[11]

Outside the large-company headlines, the pattern looks quieter. The New York Fed's regional business surveys of New York and northern New Jersey found that "only 4 percent of service firms reported laying off workers in response to AI over the past six months," while about 15% said they had hired fewer workers than they would have without AI. The researchers concluded that "existing workers are much more likely to be retrained than replaced by AI."[15]

Several large companies have adjusted course. Commonwealth Bank of Australia reversed 45 customer-service redundancies announced after it introduced an AI voice bot, saying its assessment "did not adequately consider all relevant business considerations and this error meant the roles were not redundant."[16] Klarna said in 2024 that its AI assistant was doing the work of 700 customer service agents. By 2025 its CEO acknowledged that customer service quality had fallen when cost became "a too predominant evaluation factor," and the company began recruiting human agents again.[17]

Important nuance. Some reductions appear to hold. Salesforce's CEO said support headcount went "from 9,000 heads to about 5,000" with customer satisfaction "about the same."[18] It deserves to be taken seriously, and it's also a reason to define value carefully before people decisions are made, so that quality, trust, institutional knowledge and future capacity all have a place in the value case.

MID-MARKET TRANSLATION

The evidence	Realized workforce reductions fell short of expectations, and some cuts were judged wrong or later adjusted.
What it reveals	Measuring AI's value only by headcount saved leaves out quality, customer trust and what the freed-up capacity could be used for.
At $10M–$100M	Knowledge, relationships and judgment are hard to replace. Decide the higher-value work before removing capacity.
The leadership question	**If AI gives people capacity back, what more valuable work should that capacity move toward?**

05

Progress

What evidence will show that AI improved the business, before activity metrics are treated as value?

THE BELIEF

"Time saved is value created."

Efficiency is the most common way AI value gets described, and time saved is usually how it's measured. It's easy to count and easy to communicate, far easier than better judgment or fewer errors, so it tends to become the default proof of AI value.

The belief breaks because saved time gets absorbed by lower-value work unless someone redirects it. A person saves thirty minutes and spends it on lower-priority work. A manager gains capacity and fills it with meetings. A department gets faster while nothing the customer sees improves. Leaders have to decide where the capacity goes, which outcomes should improve, how managers will guide the reinvestment, and what evidence will show the business became more effective. Without those decisions, a company can report hours saved while margin, throughput and quality stay where they were.

Gartner's May 2026 survey of 210 sales leaders found that AI saves sellers "an average of 4.8 hours per week," yet "72% of sales organizations report low reinvestment." Organizations that do reinvest are "2.2x more likely to exceed customer growth goals."[4] That's a correlation, and it's the clearest available link between redirected time and an outcome.

BCG's AI at Work 2026, with close to 12,000 respondents, found that "66% still receive limited or no guidance on what to do with the time they save," and observed that "the time that individuals save leaks out of the organization unless it is tracked and deliberately reinvested."[3]

Humlum and Vestergaard, studying chatbot use across roughly 7,000 workplaces, found that 85% of users reported reallocating time savings to other job tasks, with no measurable effect on earnings or recorded hours.[19] An ECB analysis put the median user's saving at three hours a week, about 7.7% of working time, and added that savings "only translate into higher productivity if workers turn the freed hours into extra output."[20]

Important nuance. Measured output gains do exist. Brynjolfsson and colleagues found that AI assistance raised issues resolved per hour by 14% on average across 5,179 support agents, and by 34% for novices.[21] That gain was visible because someone measured an outcome. Before time savings count as value, decide where the gain is supposed to appear and what evidence will show it.

MID-MARKET TRANSLATION

The evidence	AI saves time, but most organizations give little guidance on using it, and few reinvest it deliberately.
What it reveals	Time saved is potential value until it's connected to an outcome.
At $10M–$100M	Without that connection, the company absorbs tool cost, change effort and distraction with no gain in margin, growth, service or throughput.
The leadership question	**When AI saves time, what business outcome should improve, and how will we know?**

SIDEBAR

Measuring the input is not the same as measuring progress

As AI spending grew through 2026, several large companies reached for input controls. Uber encouraged employees to use AI coding tools "as much as possible," with usage leaderboards, then set a $1,500 monthly cap per employee per tool.[22] The Information reported that Uber's CTO said the company had "blown through its entire annual AI budget in a matter of four months."[23] Meta took down a token-consumption leaderboard that ranked about 85,000 workers; Adam Mosseri, head of Instagram, said, "It's not that hard to build a token incinerator, and that doesn't create a lot of value."[24] An OpenAI enterprise report found that "revenue per employee is not meaningfully associated with output tokens per employee or messages per active user."[25]

Samsara and Compass introduced usage caps and per-engineer budgets, and Fortune reported that some 2026 AI budgets were blown past "without producing any corresponding value to the business."[26] Business Insider reported a $2,000 monthly spending limit on an AI coding tool for some JPMorgan engineers.[27]

These controls are becoming routine. In KPMG's Q3 2026 pulse of leaders at organizations with $1 billion or more in revenue, 43% "have implemented usage or token budgets," and 74% now include cost reviews in AI approval processes, up from 61% the quarter before. Asked where AI was creating value, 37% pointed to stronger financial performance.[28] Others are moving routine work to cheaper models. On Ramp's platform, frontier models' share of business AI usage fell from a 53% peak in August 2026 to 45% in early September, as companies set lower-cost defaults.[29]

Consumption measures help control spend. They show that people use AI and say little about whether the business has become more effective. Input measures include tokens, messages, licenses and spend. Progress evidence looks different: shorter cycle times, less rework, better quality, higher throughput, lower risk, or capacity redirected to more valuable work. Both matter, and they answer different questions. Before celebrating volume, decide what value the volume is supposed to create.

SIDEBAR QUESTION

Which of our AI metrics show consumption, and which show business progress?

SELF-CHECK

Five questions before funding the next AI pilot

Before the next AI pilot receives budget, ask your leadership team to answer these five questions.

Mark each answer clear enough to fund, partially clear or unclear. If two or more are partially clear or unclear, the next pilot may succeed on its own and still leave the leadership team no clearer about what to fund next.

DIRECTION ☐ Clear enough to fund ☐ Partially clear ☐ Unclear

What business direction are we asking AI to support before we evaluate tools, platforms, or partners?

DECISIONS ☐ Clear enough to fund ☐ Partially clear ☐ Unclear

What criteria determine which AI opportunities deserve budget, capacity, and executive attention?

WORK ☐ Clear enough to fund ☐ Partially clear ☐ Unclear

Which workflows, decisions, handoffs, and review points need to change before AI can improve the business?

PEOPLE ☐ Clear enough to fund ☐ Partially clear ☐ Unclear

If AI gives people capacity back, what more valuable work should that capacity move toward?

PROGRESS ☐ Clear enough to fund ☐ Partially clear ☐ Unclear

When AI saves time, lowers cost, improves quality, or increases output, what business outcome should change, and how will we know?

A leadership team doesn't need perfect answers before every experiment. It needs enough clarity to know which decision the experiment is meant to improve.

Before you fund the next AI pilot, align the leadership conditions around it

The AI Investment Alignment Workshop helps leadership teams examine the five conditions that shape whether AI activity becomes business progress. In this half-day working session, JR Key Advisory facilitates a structured discussion of each one. Before the session, each leader completes the free Executive AI Alignment Diagnostic, so the team arrives with its views already compared by condition.

Direction: what AI should help the business become better at.

Decisions: which opportunities deserve budget, capacity and executive attention.

Work: which workflows, handoffs and review points need to change.

People: where the capacity AI returns should go.

Progress: what evidence will show the business improved.

> Designed for three to eight leaders who decide on and oversee AI work. Delivered virtually or on-site. Investment: $3,500, credited in full toward the AI Strategy and Transformation Readiness Assessment or the Executive AI Strategy Intensive if the team continues within the agreed period. Participants leave with the AI Investment Alignment Snapshot: where the team aligns and diverges on each condition, the conditions it can already build on, and a recommended next step.
>
> **Explore the AI Investment Alignment Workshop: jrkeyadvisory.com/services.html#alignment-workshop**

If your leadership team can't yet answer the five self-check questions clearly, the most useful next investment may be an alignment conversation, held before the next pilot is funded. None of those five questions can be answered by a tool. They're answered by leaders deciding together, which is why, even in an age of AI, human collaboration is still the ultimate competitive advantage.

About this paper

JR Key Advisory, based in Portland, Oregon, is an independent AI strategy advisory for regulated, trust-sensitive and knowledge-intensive mid-market organizations. It helps leadership teams clarify what AI should accomplish, where the organization is ready, which initiatives deserve priority, and whether implementation is producing the progress executives intended.

This paper draws on publicly reported research, surveys and company examples.

Endnotes

1. Deloitte, State of AI in the Enterprise 2026, 21 Jan 2026; 3,235 leaders in 24 countries.
2. KPMG, Global AI Pulse Q2 2026; n=2,145.
3. BCG, AI at Work 2026, Jun 2026; close to 12,000 respondents.
4. Gartner, press release, 19 May 2026; 210 chief sales officers and senior sales leaders. Correlational.
5. Gartner conference session, "Don't Let Vendors Run Your AI Strategy," Denis Torii, 22 Sep 2026 (session description).
6. Ryseff, De Bruhl and Newberry, The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed, RAND Corporation, RR-A2680-1, 13 Aug 2024.
7. Kim, Kim and Koning, "Mapping AI into Production: A Field Experiment on Firm Performance," INSEAD Working Paper 2026/20/STR; SSRN 6513481. Randomized field experiment with 515 high-growth firms; working paper.
8. MIT NANDA, as reported by Fortune, Aug 2025.
9. Gartner, press release, 1 Sep 2026; 1,303 respondents at organizations with revenue of $50M or more; fielded Jan–Apr 2026.
10. Ferrazzi, Smith and Roberts, op-ed, Fortune, 19 Mar 2026.
11. McKinsey, The state of AI in 2026: On the road to ROI, 25 Aug 2026; 1,719 respondents; workforce figures based on 1,521 AI-using organizations; high performers defined by reported outcomes.
12. PwC, AI Performance Study, 13 Apr 2026; 1,217 senior executives.
13. Allianz, Project Nemo, allianz.com, 3 Nov 2025. Company-reported results.
14. Intelligent Insurer, 6 Oct 2026 (Tina Baacke, Swiss Re Corporate Solutions).
15. Abel, Deitz, Emanuel and Montalbano, "Businesses Are Using AI to Transform Work, Not Cut Jobs," Liberty Street Economics, Federal Reserve Bank of New York, 1 Sep 2026. Regional business surveys; self-reported.
16. ACS Information Age and The Canberra Times, 21 Aug 2025.
17. CNBC, 14 May 2025; Fortune, 9 May 2025, reporting Bloomberg.
18. Salesforce CEO Marc Benioff, remarks as reported Sep 2025.
19. Humlum and Vestergaard, NBER Working Paper w33777, revised Mar 2026. Self-reported survey data.
20. European Central Bank, The ECB Blog, 26 Aug 2026. Self-reported.
21. Brynjolfsson, Li and Raymond, "Generative AI at Work," Quarterly Journal of Economics 140(2), 2025 (NBER Working Paper w31161).
22. TechCrunch, 2 Jun 2026.
23. The Information, as reported by TechCrunch.
24. Fortune, 12 May 2026; TechCrunch, 14 Jul 2026.
25. OpenAI enterprise report, as reported by Fortune, 13 Aug 2026. Vendor-authored.
26. Fortune, 12 Aug 2026.
27. Business Insider, as reported by Newsquawk, 17 Sep 2026.
28. KPMG, Q3 2026 AI Pulse, 24 Sep 2026; 314 US leaders at organizations with $1B or more in revenue. Self-reported.
29. Ramp, Ramp AI Index, Ramp Economics Lab, 9 Sep 2026; Ramp customer spend data.

JR Key Advisory

AI activity is not AI progress.