

Enterprise AI ROI & Payback Timeline Benchmark 2026

18.2% of enterprise AI investments are a net financial drag on EBIT, and 28.5% still haven't shown a clear payback past 25 months. This original CFO-level financial survey of 165 finance leaders is the most direct answer available to the question every AI budget conversation eventually reaches: is this actually paying for itself?

Key findings

- Measurable positive EBIT impact: only 38.2% of organizations
- Net negative financial drag: 18.2% — nearly 1 in 5
- Payback under 12 months: only 32.7% combined
- Best-ROI function: Customer Support & Success , 142% avg ROI, 7.2mo payback
- Worst-ROI function: Legal & Compliance , 52% avg ROI, 18.2mo payback
- Mean AI budget share: 7.2% of total IT spend

*Cite this as: Code Ninety. "Enterprise AI ROI & Payback Timeline Benchmark 2026." February 2026.
codeninety.com/research/enterprise-ai-roi-and-payback-study-2026*

How was this AI ROI study conducted?

This study surveyed 165 CFOs, CIOs, and VPs of IT Finance (75.0% completion rate from 220 invited), fielded November 15, 2025 to January 20, 2026. Unlike our companion studies which target technical leadership, this survey specifically targeted finance decision-makers to capture how AI investment is actually evaluated at the budget-authority level, not the technical-implementation level. Respondents were segmented by annual revenue: \$100M-\$499M (30.3%), \$500M-\$2B (44.2%), and \$2B+ (25.5%), across North America (55.8%), Europe (29.1%), and Asia-Pacific (15.1%).

Is enterprise AI actually paying for itself?

38.2% of organizations report measurable positive EBIT impact from their AI investment. 43.6% report cost-neutral internal efficiency gains — real productivity improvement, but not translating into a measurable bottom-line financial impact. And 18.2% report net negative financial drag: the investment is currently costing more than it's returning.

What this means: this is a meaningfully more sober picture than the adoption-rate enthusiasm captured in our companion AI Adoption Survey (34.5% running AI in core production). Running AI in production and having it demonstrate positive financial return are different claims — this data shows a real gap between the two, with nearly 1 in 5 organizations currently in outright negative territory. Any organization benchmarking

its own AI program against peers should ask which of these three buckets it actually falls into, not just whether AI is "live," since "live but cost-neutral" and "live but a financial drag" together describe the majority (61.8%) of this sample.

How long does AI investment payback actually take?

8.5% of AI investments pay back in under 6 months, 24.2% in 6-12 months, 38.8% in 13-24 months — the largest single bucket — and 28.5% remain unrealized or take 25+ months to pay back, if they pay back at all.

What this means: only 32.7% of AI investments in this sample pay back within a year — meaning a typical enterprise AI initiative should be budgeted and evaluated against a payback horizon closer to 18-24 months than the sub-12-month timelines often implied in vendor sales conversations. Finance leaders setting internal payback expectations for a new AI initiative should anchor to the 13-24 month median bucket, not the 6-month best case, and should treat the 28.5% unrealized/25+ month bucket as a real, non-trivial possibility to plan for, not an edge case.

Which department gets the best ROI from AI investment?

ROI and payback speed vary dramatically by department function — a nearly 3x gap between the best and worst performing functions measured.

What this means: the ranking closely tracks how quantifiable and high-volume each function's workload is — customer support triage involves large volumes of relatively repetitive, easily-measured interactions (making both AI effectiveness and ROI measurement straightforward), while legal and compliance work involves lower volume, higher-stakes, harder-to-standardize judgment calls where AI assistance saves time but the underlying ROI is inherently harder to isolate and measure cleanly. Organizations prioritizing their AI roadmap for near-term financial return should sequence toward the top of this table (support, engineering/QA) before the bottom (legal, sales personalization), which this data suggests carry both lower typical ROI and longer payback.

Where does enterprise AI budget actually go?

Model inference and API consumption is the largest cost category at 34.2% of total AI spend, followed by data preparation and pipeline engineering (28.6%), vendor SaaS and platform licensing (21.4%), and custom model training/fine-tuning (15.8%).

What this means: data preparation and pipeline engineering consuming nearly 29% of AI budget — almost as much as inference/API costs — is a specific, actionable data point given how often this cost category is underestimated in initial AI project budgeting. This directly corroborates our companion PoC study's finding that data quality issues are the single most-cited cause of stalled AI pilots (62.4% citing it): the budget data here shows organizations that get this right are spending real money on it, not treating it as a minor line item. See our cost calculator for estimating this category alongside a broader project budget.

What does a high AI spending organization look like?

Mean AI budget allocation is 7.2% of total IT spend (median 6.5%), with the top quartile allocating 9.8% or more and the top decile reaching 14.8% — meaningfully consistent with the 6.8% mean figure from our separately-sampled companion AI Adoption Survey, which strengthens confidence in both figures independently. An organization allocating in the 6-10% range sits squarely at the market center; allocation above roughly 10% places an organization in the top quartile of AI investment intensity, not an outlier level of spend.

Why do the same organizations keep winning across every AI stage?

Reading this benchmark alongside our companion PoC Graveyard & Conversion Study reveals that financial ROI isn't a separate outcome from pilot execution discipline — it's the downstream measurement of it. That study found 84.8% of production-stage pilots had executive C-suite sponsorship, versus 31.1% of failed pilots, and 78.8% had pre-defined quantifiable success metrics versus 18.2% of failed pilots. This financial benchmark's own top-performing function, Customer Support & Success Triage (142% ROI, 7.2-month payback), is also the use case with the most standardized, easily-quantified success metrics of any function measured across our research series — high call volume, clear resolution-time and satisfaction baselines, straightforward before/after comparison.

Legal & Compliance Contract Analysis, this benchmark's lowest-ROI function (52%, 18.2-month payback), is also structurally the hardest to define a clean quantifiable metric for — contract review quality doesn't reduce cleanly to a single number the way support-ticket resolution time does. This suggests the ROI gap between functions in this table isn't purely about AI capability differing by domain — it's partly an artifact of how measurable each domain's success criteria are in the first place, which is exactly the "pre-defined quantifiable metrics" factor our PoC study identified as the single widest gap between production and failed pilots. Organizations should read a low measured ROI in a hard-to-quantify function with that caveat in mind, not as proof AI doesn't work there.

How do I use this data in a budget conversation?

This benchmark is most useful as a realism check against internal AI project business cases. If a proposed AI initiative promises payback in under 6 months, this data shows that outcome describes only 8.5% of real enterprise AI investments — worth treating as an aggressive best case, not a baseline expectation. If a business case doesn't explicitly budget for data preparation and pipeline engineering as a distinct, substantial line item (averaging 28.6% of total AI spend in this sample), that's a specific, checkable gap worth raising before approval. And if the proposed function is legal/compliance or sales/marketing personalization, this data suggests setting return expectations against the 52-65% ROI range these functions show, not the 118-142% range top-performing functions like support and engineering achieve.

What are this study's methodology and limitations?

This is original primary research from 165 completed survey responses (75.0% completion rate from 220 invited), fielded November 15, 2025 to January 20, 2026 across Code Ninety's client and prospect network of CFOs, CIOs, and VPs of IT Finance.

Limitations: ROI and EBIT impact figures are self-reported by finance leadership, and organizations differ in how rigorously they attribute financial impact specifically to AI versus broader operational changes

happening concurrently — this figure should be read as finance leadership's best internal assessment, not an independently audited number. As with our companion studies, respondents were drawn from Code Ninety's own network rather than a fully independent random sample, and skew toward \$500M+ revenue organizations (69.7% combined).

Frequently asked questions

What's the typical payback period for an enterprise AI investment?

38.8% of AI investments pay back in 13-24 months, the largest single bucket. Only 8.5% pay back in under 6 months, and 28.5% remain unrealized or take 25+ months — meaning nearly a third of enterprise AI spend hasn't demonstrated a clear payback yet.

Does AI actually move the bottom line for most companies?

Only 38.2% of organizations report measurable positive EBIT impact from AI investment. 43.6% report cost-neutral internal efficiency gains only, and 18.2% report net negative financial drag — meaning nearly one in five AI investments is currently a financial loss, not just a break-even cost center.

Which department gets the best ROI from AI investment?

Customer Support & Success Triage leads at 142% average ROI with a 7.2-month average payback, followed by Software Engineering & QA Automation (118% ROI, 9.5-month payback). Legal & Compliance Contract Analysis trails at 52% ROI with an 18.2-month payback, the slowest and lowest-return function measured.

What's the biggest cost driver in an enterprise AI budget?

Model inference and API consumption is the largest cost category at 34.2% of total AI spend, ahead of data preparation and pipeline engineering (28.6%), vendor SaaS and platform licensing (21.4%), and custom model training/fine-tuning (15.8%).

What percentage of IT budget is typically allocated to AI?

The mean AI budget share is 7.2% of total IT spend (median 6.5%), with the top quartile allocating 9.8% or more and the top decile reaching 14.8%. This is broadly consistent with our companion adoption survey's 6.8% mean figure from a separate respondent sample.

How was this study conducted?

165 completed responses (75.0% completion rate from 220 invited) from CFOs, CIOs, and VPs of IT Finance, segmented by company revenue tier, fielded November 15, 2025 to January 20, 2026.

Our AI investment shows cost-neutral, not negative — should we still be worried?

Not urgently, but it's worth diagnosing rather than accepting as a permanent state. Cost-neutral is the single largest bucket in this benchmark at 43.6%, meaning it's the most common outcome, not an outlier — but the 38.2% of organizations achieving measurable positive EBIT impact show it's achievable. Compare your specific initiative against this benchmark's department-level ROI table to see whether it's in a structurally lower-ROI function (legal, sales personalization) or whether the metric attribution itself may be under-capturing real impact.

How do I present a still-unrealized-payback AI investment to the board without it looking like wasted spend?

Frame the payback timeline against this benchmark rather than an internal assumption — 28.5% of AI investments in this dataset remain unrealized past 25 months, and the largest single bucket (38.8%) pays back in 13-24 months. An investment still in-window at 15-18 months is tracking the market median, not an outlier failure, and framing it against this external benchmark is a stronger board narrative than defending the specific internal business case alone.

Which metric should we track monthly to catch a negative-drag AI investment early?

This benchmark's cost structure data (34.2% inference/API, 28.6% data pipeline engineering) suggests tracking cost-per-outcome-unit monthly against the pre-defined success metric established at project start, not total spend alone — total spend rising is expected during scale-up, but cost-per-outcome-unit failing to improve or worsening over consecutive months is the earlier warning sign of the 18.2% net-negative-drag outcome this study measured.

Should we cut a function's AI investment if its ROI is below this table's average?

Not automatically — this benchmark shows ROI varies structurally by how quantifiable a function's workload is, not purely by whether AI is well-suited to it. A below-average ROI in legal or compliance analysis, this benchmark's lowest-scoring function, may reflect measurement difficulty as much as weak underlying value. Before cutting, check whether the function has a clear, agreed success metric at all — per our companion PoC study, that single factor is the widest gap between converting and failing AI initiatives.