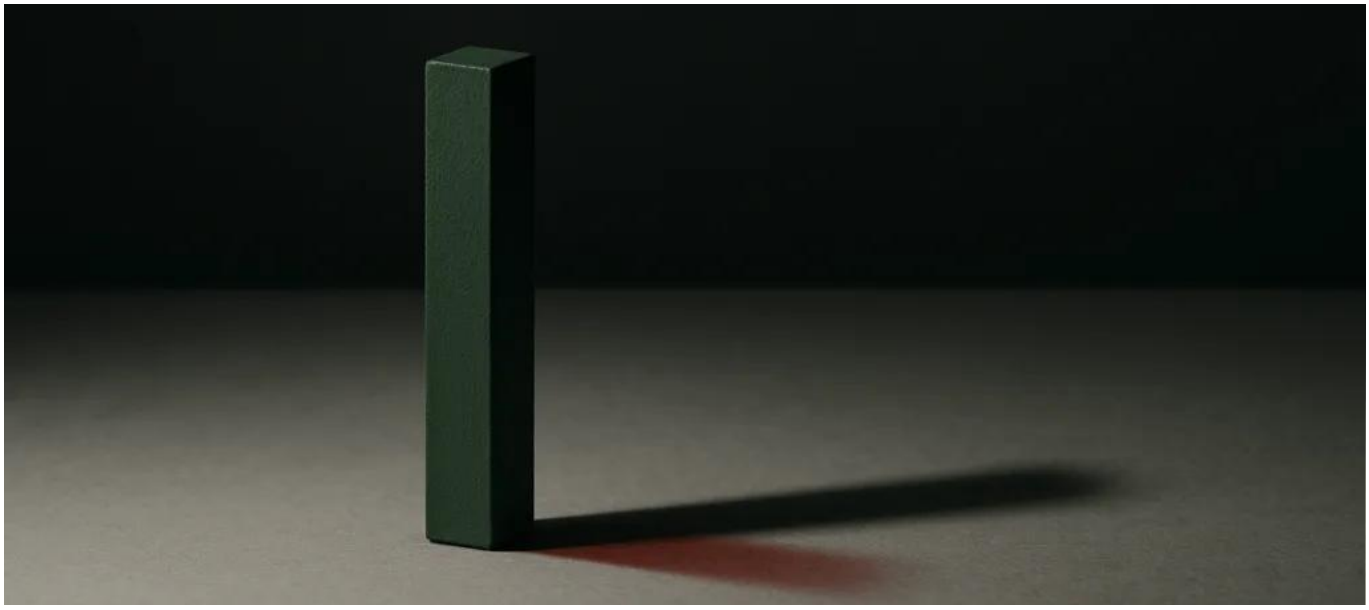

Enterprise AI ROI: 5 Numbers Everyone Quotes Wrong, From “95% Fail” to McKinsey’s 37%

Six percent of McKinsey’s 1,719 respondents qualified as AI high performers. That single figure is the entire source of the “94% of enterprises see nothing” headline that has been circulating since August.

The number worth quoting is 63%. In The state of AI in 2026: On the road to ROI (published August 25, 2026, fieldwork May 4 to June 8), 37% of respondents attributed any positive EBIT impact to AI, down marginally from 39% in 2025. Underneath that: 6% high performers, roughly 31% reporting sub-threshold impact, and 63% reporting no measurable enterprise earnings impact at all.

Three outlets read the same dataset and published “on the road to ROI” (The Register), “94% of enterprises” (TechTimes), and “the ROI didn’t come” (TechTarget). None of them lied. They each picked a different true cut of the same table, which is why the number you

carry into a board meeting depends mostly on which tab you had open. The cost of getting that wrong is not pedantic: a 94% figure argues for cancelling the programme, and a 63% figure argues for building an attribution model.

The 94% is the complement of the 6% high performers

THE CLAIM	THE REALITY
94% of enterprises get nothing from their AI spending.	94% is the arithmetic complement of the 6% who cleared McKinsey’s high-performer bar: at least 5% of EBIT attributed to AI plus significant described organizational impact. The share reporting no measurable earnings impact is 63% (McKinsey / TechTimes, August 2026).

The 31-point gap between 94% and 63% is not a rounding dispute. It separates “almost nobody gets anything” from “roughly a third get something real but not transformative.” Those two sentences support different budget decisions.

Myth two is the MIT one. The claim is that MIT found 95% of AI pilots fail. What Project NANDA actually found is that roughly 5% of integrated pilots generated significant value against $30 to $40 billion invested in enterprise GenAI. TechTarget’s September 25, 2026 analysis says the result was “commonly simplified, too broadly, into the claim that 95% of AI fails.” “Did not generate significant value” and “failed” are different categories. A pilot that works, ships, and produces a modest efficiency gain nobody bothered to quantify sits in the 95% bucket. So does a pilot abandoned in week three. Lumping them together tells you nothing you can act on.

THE CLAIM	THE REALITY
40% of companies have scaled AI agents.	40% is the figure for organizations above $1 billion in annual revenue, up from 27%. It is not the all-company number. This is one of the most common misquotes from the 2026 survey (McKinsey, August 25, 2026).

Deployment rose six points while EBIT attribution fell two

Deployment and attribution get measured by different people, on different timelines, with different incentives. Scaling across the enterprise went from 38% to 44% year over year. Nearly nine in ten respondents (1,521 of 1,719) reported regular AI use in at least one business function. The EBIT line went from 39% to 37%.

My read: the flat EBIT number is only partly a performance story. It is also a measurement story, and the German spotlight makes that visible in a way the global figure does not.

43%
German respondents who cannot determine AI’s contribution to EBIT at all
McKinsey Germany spotlight, Sep 8 2026, 93 weighted respondents

33%
German respondents reporting positive EBIT contribution
McKinsey Germany, Sep 8 2026, vs 37% globally

63%
global respondents reporting no measurable enterprise earnings impact
McKinsey / TechTimes, Aug 2026

Read those two German numbers together. 33% say yes, 43% say they cannot tell. That leaves at most 24% who looked and found nothing. The dominant category in Germany is the absence of an instrument. When 43% of a sample has no attribution mechanism, a survey about EBIT impact is measuring the state of finance reporting as much as the state of AI.

The EU size gradient matters more than the 20% headline

Eurostat put EU enterprise AI use at 20.0% in 2025, up from 13.5% in 2024. That 6.5 point jump is the headline and the least interesting part of the release. The size gradient is the finding: 55.03% of large enterprises used AI in 2025, against 30.36% of medium and 17.00% of small ones.

A large European enterprise is more than three times as likely to use AI as a small one. That spread has an obvious reading, large firms have the capital and the data engineering, and a less comfortable one. The McKinsey EBIT survey skews toward exactly the organizations where AI is already everywhere and still not showing up in earnings. The 40% agent-scaling figure for billion-dollar-plus firms sits in the same place. Deployment concentrates at the

top, and the top is where attribution is hardest, because a $5 billion P&L absorbs a lot of efficiency before it registers.

✓ **The fix**
Before the next AI budget cycle, write down the specific line item that is supposed to change, who owns it, and what it reads today. If nobody can name the line, you are about to join the 43% who cannot determine contribution, and no amount of model quality will rescue that.

A related pattern shows up in procurement rather than earnings. I wrote earlier about 32% of companies killing a software purchase because coding agents could build the thing instead. That saving is real and it lands in a vendor budget, not in an AI programme’s ROI column. Whoever runs the AI steering committee rarely gets credit for a SaaS contract that was never signed.

Measurement failure looks exactly like value failure

MY TAKE

I think the flat 37% is roughly half genuine disappointment and half attribution failure, and I cannot prove the split from this data. What I can prove is the shape: 63% report no measurable impact, 43% of the German subset report no ability to measure, and deployment rose 6 points while attribution fell 2. Enterprise AI’s earnings contribution is currently unobservable at most firms, which is a different problem from being zero, and it needs a finance fix before it needs a model fix. These myths persist because a complement is easier to tweet than a breakdown.

The same confusion between perception and instrument shows up at the individual level. METR found developers believed they were faster with AI while the stopwatch said they were 19% slower. At company scale the error runs the other way. Executives underclaim, because EBIT is noisy, lagging and heavily attributed, and nobody wants to be the person who told the board that AI added margin and then had to defend the arithmetic.

Five numbers worth fixing in your own deck. 94% is a complement, not a finding. 63% is the real bad number. 95% of pilots did not fail; 95% did not produce measured significant value

against the $30 to $40 billion spent. 40% agent scaling applies only above $1 billion in revenue. And 37% against 39% is flat, which makes the 2026 survey’s most defensible conclusion this: a year of scaling moved deployment and left attribution where it was.

WRITTEN BY

Artur Markus

Software engineer in Basel building AI infrastructure and agentic systems, with a background in pharmaceutical automation, GxP compliance, and validation. I help teams turn AI from a chatbot into a reliable operational layer.

Book a meeting →