

How to calculate *the ROI* of AI

The metrics framework we use to defend an AI project to leadership or a board. With the formulas, a worked example, a calculator and the one-page template to present it.

Online version, with calculator: adandesign.com/en/ai-roi-guide/ · Adan Design,
Mérida, Mexico · October 2026

◆ IF YOU ONLY HAVE TWO MINUTES

- A board doesn't approve "productivity". It approves hours that get reassigned, costs that disappear, errors that drop or revenue that grows, measured against a baseline.
- Hours saved are only worth money if someone decides what they go to. Count only the reassigned share.
- The cost of AI isn't the license. It's license, implementation, training, human review and maintenance.
- A 30-day pilot with one task, one metric and a baseline is worth more than any spreadsheet projection.
- Bring two scenarios: yours and the cautious one. The person who shows their own pessimistic version is the one who gets believed.

01

Why they don't believe you

Almost every AI proposal we've seen reach a board fails in the same place. It talks about productivity.

"We'll be 30% more productive" means nothing to the person signing the budget, because they don't know what changes on Monday: which task, whose, how long it takes today, how long it will take afterwards and what happens to the time left over. And if you don't know either, it shows.

The three mistakes that come up most:

1. **No baseline.** Nobody measured how long the task takes today, so any improvement is an opinion.
2. **Hours saved counted as money.** If a salesperson saves five hours a week and nobody decides what to do with them, the company saved nothing. It paid the same.
3. **The license as the only cost.** Twenty dollars a month per seat sounds like nothing. What costs is having someone set the tool up, train the team, review what comes out and keep it alive when the model changes, which happens every few months.

Our position, so there are no surprises later: we don't project headcount cuts as return. They almost never happen that way, and a project that starts with that promise sabotages itself, because the people who have to adopt it stop helping. If someone promises to cut half your team with a chatbot, be skeptical.

02

The four drawers of return

Every return from an AI project falls into one of four drawers. If it doesn't fit in any of them, it isn't return. It's a feeling.

A

Hours recovered

Time a task stops consuming. Measured in minutes per task times tasks per week. It's worth money only for the share that gets reassigned to something of value: more customers served, more proposals sent, fewer overtime hours paid.

B

Direct cost avoided

Money you stop paying: an outside vendor, overtime, a tool the AI replaces, courier, printing. It's the easiest drawer to defend because it shows up in the books without anyone having to believe you.

C

Errors and risk

Rework, returns, penalties, botched quotes, late replies that lose customers. Measured as a rate (errors per 100 tasks) and by what each error costs.

D

Revenue

More quotes sent, more leads answered on time, content that wasn't being produced before. It only counts if you already know your conversion rate. If you don't, it goes in as a hypothesis to validate in the pilot, not as a number in the projection.

One piece of advice that comes from watching it repeat: pick one or two drawers per project. A proposal that promises all four at once looks like a wish list, and the board treats it like one.

03

The costs everyone forgets

On the other side of the formula goes everything the project costs, and this is where most proposals fall short. Here's the full list.

COST	WHAT IT INCLUDES	WHEN YOU PAY IT
Licenses	Tools billed per seat or per use (tokens, API calls).	Monthly
Implementation	Hours of whoever sets up, connects to your systems and tests, whether in-house or a vendor.	Once
Training	The sessions, people's time in them and the weeks when they're still slower than before.	Once per person
Human review	Someone checks what comes out before it reaches a customer. Always. Count their minutes per task.	Every task
Maintenance	Adjusting prompts, flows and connections when the model changes or your process changes. Our rule: set aside 10 to 20% of the implementation cost every year.	Yearly
Data and security	Cleaning the information the tool will use and writing down what can go in and what can't.	Once, with reviews

Human review is the cost most often left out and the one we care about most. It's a design decision: what comes out of a model gets checked by a person before it reaches the customer, and those minutes go in the count. When you add them, the proposal loses some shine and gains credibility. It comes out ahead.

04

The formulas

Four. You don't need more.

Monthly value of hours recovered = people × hours saved per person per week × cost per hour × 4.33 × % reassigned

Cost per hour is the fully loaded cost (salary, benefits and what it takes to keep that person) divided by paid hours. Ask finance for it instead of estimating: it's the number they'll question most.

Net return, year 1 = benefits for the year (the drawers you measured) – costs for the year (upfront cost + 12 × monthly cost)

Benefits only from the drawers you measured in the pilot. Hypotheses stay out until they have data.

ROI = net return ÷ costs for the year × 100

How much each dollar invested earns, as a percentage. An ROI of 100% means the project returned twice what it cost.

Payback in months = upfront cost ÷ (monthly benefit – monthly cost)

When it pays for itself. For a board this usually weighs more than ROI: a project that pays back in four months gets approved even with a modest return. One with a spectacular return over three years, almost never.

05

A 30-day pilot

Before the pilot, your projection is a guess. This is how you turn it into data.

WEEK 0

Baseline

Pick a repetitive task with a clear owner that happens several times a week. Measure two weeks without AI: minutes per task, tasks per week, errors. A three-field form is enough. If the data already lives in a system, even better.

WEEKS 1 TO 4

Pilot

Three to five people, the same task, with the tool and with written rules of use: what information goes in, who reviews, what to do when the AI gets it wrong. They measure the same things as in week 0. Also note how many AI outputs were approved without changes; that figure is your quality indicator and tells you what review costs.

Count and project

Compare against the baseline using the real costs the pilot had, not the budgeted ones. And project cautiously: for year 1 we use 60 to 70% of the pilot's result, because the pilot is run by the most motivated people and everyone else adopts more slowly.

An example from home, not about AI but about measuring. The hackathon-based recruiting model at Genius Arena, where Adan has run the Talent Hackathon since 2016, evaluates dozens of candidates in one week and cuts recruiting time by 80% compared with the traditional process. That number exists because the old process was measured before it was changed. Without that first measurement it would be an anecdote, and anecdotes don't get past boards.

A worked example

The figures are illustrative, in Mexican pesos, so you can see the method. They aren't from a client.

A distributor with four salespeople. Each puts together about 30 quotes a week and spends 50 minutes on each one, between looking up prices, writing and formatting. With an assistant set up with their catalog and their rules, and with the salesperson reviewing before sending, a quote drops to 20 minutes. Review time is already inside those 20.

WHAT WAS MEASURED	BASELINE	WITH THE ASSISTANT
Minutes per quote	50	20
Quotes per week (4 salespeople)	120	120
Hours recovered per week	–	60
Loaded cost per salesperson hour	\$250	\$250
Share reassigned (to visits and follow-up)	–	50%
Useful value per month (60 × 250 × 50% × 4.33)	–	\$32,475

YEAR 1 COSTS	AMOUNT
Implementation: set up the assistant with the catalog, the rules and the tests	\$60,000 once
Training for the four salespeople and the manager	\$25,000 once
Licenses, 5 seats	\$2,000 per month
Maintenance	\$10,000 per year
Year 1 total	\$119,000

RESULT	YOUR SCENARIO	CAUTIOUS SCENARIO (65%)
Year 1 benefit	\$389,700	\$253,300
Net return	\$270,700	\$134,300
ROI	227%	113%
Payback	3 months	5 months

Notice what isn't in the table. If quoting in 20 minutes raises the close rate because the customer gets the proposal the same day, that's drawer D. It gets measured in the pilot

and, if confirmed, goes into the second-half projection. Not before. And even in the cautious scenario, the project pays for itself in five months. That's the sentence you want to say in the room.

07

The metrics for the board

Eight. You need the first four in any project; the rest depend on the drawer you picked.

METRIC	HOW TO MEASURE IT	WHAT IT'S FOR
Minutes per task	Stopwatch or system timestamp, on every task, two weeks before and four during.	Drawer A. Without this, no guide will help.
Hours reassigned, and to what	One weekly question to the owner: what did the time left over go to?	Turns hours into money. No answer means the saving is zero.
Outputs approved without changes	Out of every 10 AI results, how many went out as is.	Quality of the system and size of the review cost.
Adoption	People using the tool at least three times a week, out of those who should.	An unused tool has zero return. It's the first warning sign.
Tasks per person per week	Simple count.	Shows whether recovered time turns into more work done.
Errors or rework	Errors per 100 tasks, before and during.	Drawer C.
Customer response time	Hours between request and delivery.	Drawer C, and D if you have conversion data.
Cost per task	Month's costs divided by month's tasks.	The number finance understands best.

What not to promise

- **Headcount cuts.** We said it above and we'll say it again, because it's the promise that kills the most projects before they start.
- **"10x" or "double the productivity".** If it doesn't come from your own measurement, the board takes it for what it is: a headline.
- **Savings from month one.** For the first two months people are slower, not faster. Put it in the projection and nobody is surprised.
- **Return from a tool nobody uses.** If adoption stalls at two people, there's nothing to count. That's why we measure it from week one.
- **Results without human review.** Promising the system "no longer needs anyone to check it" is the fastest way to lose the board's trust the day it gets something wrong. And it will.

Calculator

Enter your numbers and it gives you the monthly value, year 1 return, ROI and months to payback, with the cautious scenario next to each. Works in any currency. It stores nothing and sends nothing.

One-page template

What you take to the board fits on one sheet. If it needs more, a measurement is missing. Copy it, fill in the brackets and delete what doesn't apply.

Project: [name, in six words]

Task: [which task, whose, how many times a week]

Baseline (measured from [date] to [date]):

[minutes per task] · [tasks per week] · [errors per 100]

Pilot ([dates], [n] people):

[minutes per task] · [% of outputs approved without changes] · [adoption: n of n people]

Year 1 return: [benefit] – [cost] = [net]. ROI [x%]. Pays back in [n] months.

Cautious scenario (65%): ROI [x%]. Pays back in [n] months.

Costs: implementation [amount] · training [amount] · licenses [amount] per month ·
maintenance [amount] per year · human review [n] minutes per task.

Risks and how we cover them: [what data stays out of the tool] · [who reviews each output] ·
[what we do if the model changes].

What I'm asking for: [amount] and [n] weeks to [what].

If approved: we start on [date] and review results on [date].