bosio digital

GUIDE
2026 EDITION

The *human* domain

Why the rollout stalled, and what it cost you that nobody measured.

TIME NEEDED
Thirty minutes, then one meeting

YOU LEAVE WITH
A scored map, and five controls to install

BY
Laura Pretsch
bosio.digital

BEFORE YOU START

Nobody took your judgment. *You handed it over.*

One reasonable decision at a time, and every time you did it you were rewarded for it.

An executive reads an AI-generated analysis of a business they have run for a decade. Something in it does not match what they know. They approve it anyway. Not because the machine was persuasive, but because they could not produce a defensible reason fast enough, and a view you cannot put on a slide has been the wrong kind of view for a long time now.

That is not a technology failure. It is a trained response, and the ground was prepared carefully, by people who meant well. A lot of them were consultants. Some of them were us.

THE WHOLE INDUSTRY IS LOOKING SOMEWHERE ELSE

Hallucination rates. Benchmarks. Evaluation harnesses. Guardrails. Every bit of it aimed at making the output more trustworthy. None of it aimed at the person who already suspected the output was wrong and signed it anyway. You cannot fix that with a better model.

WHY THIS MATTERS TO THE ROLLOUT THAT STALLED

92% of data and AI leaders say the barrier is people rather than technology. If your team is not using what you bought, the diagnosis is almost never the tool. This guide is the other half of that sentence: what the work does to the person, and what to put back.

THE DISTINCTION EVERYTHING RESTS ON

Checking and judging are not the same job.

Your team is checking more AI output than it has ever checked in its life. That is true, measurable, and not the good news it sounds like, because "checking" is one word doing the work of two jobs.

Detection

CATCHING THIS ERROR, IN THIS OUTPUT, NOW. LOCAL, TACTICAL, GETS FASTER WITH PRACTICE. THE ONLY ONE ANYBODY COUNTS

Diagnosis

SEEING WHY THIS KIND OF ERROR KEEPS HAPPENING. CANNOT BE DONE FROM INSIDE THE FLOW, BECAUSE THE POINT OF IT IS STANDING SOMEWHERE ELSE

Notice which one your company measures.

WHAT FULL VERIFICATION STILL MISSED

In January a technology columnist built himself a verification system: an agentic workflow and an AI critic running a thirty-five page rubric for accuracy, grammar, structure and sourcing. He pointed it at a draft about Apple, Google and Siri. It verified every claim in the piece. It also got the most important one backwards: the draft said Apple pays Google roughly twenty billion dollars a year, and Google pays Apple.

His own explanation is the sharpest sentence written about AI verification this year. The system had confirmed that "twenty billion", "Google", "Apple" and "search" all appeared together in the sources. It never asked who was paying whom. These tools, he wrote, excel at confirming that components exist in sources and struggle to confirm that relationships between components are correctly represented.

What caught it was a person who knew about the deal. Not somebody checking faster. Somebody holding the whole picture.

THIS IS NOT A NEW PROBLEM, AND IT IS MEASURABLE

Hand-classifying 570 code review comments, researchers found 65 local logic errors and 6 design-level problems. Checking frequency and checking depth are different variables. Most organizations optimize only the first.

THE INTERVAL

Two pauses, and only one can be installed.

1 Triggered by the decision

It sits in front of anything consequential, it is scheduled, and it is a control. Nobody signs until somebody has produced an objection. You can write that into a process this quarter and it will hold.

2 Triggered by a signal

The work is coming apart. The same problem keeps arriving in slightly different clothes. You have run at it four times and it is no sharper than it was on the first pass. That is when you stop, to find out why the thing will not resolve, which is a different question from the one you have been asking.

You cannot schedule the second one. Nobody knows in advance which Tuesday it lands on.

THE WHOLE ARGUMENT IN ONE TURN

Noticing that a problem has stopped clarifying is itself an act of judgment. The capacity you need to recognize you need the second pause is the capacity the second pause was going to restore.

That is circular, and it is also exactly what is happening. It is why this degrades quietly instead of announcing itself, and why nothing in your reporting will catch it. A team that has lost the second pause does not file a ticket. It keeps going, productively, into the same wall.

The first pause is what a process can give you. The second is what the first one is buying time to rebuild.

THE GAP DID NOT DISAPPEAR. IT GOT OCCUPIED

An eight-month field study inside a two-hundred-person company found workers filling exactly the moments that used to be empty. Prompting over lunch, in meetings, while a file loaded. A meta-analysis pooling 114 effect sizes found the incubation effect is real with one hard condition: fill the gap with a cognitively demanding task and most of the benefit disappears. An undesigned pause is not the intervention. A protected one is.

Four findings, arriving from four directions *at once.*

WHAT WAS STUDIED	WHAT IT ESTABLISHES
Sycophancy Science, March 2026, N=1,604	A single conversation with an AI that validates you measurably increases self-righteousness and reduces willingness to repair a conflict
Emotion vectors Anthropic, April 2026	The agreement drive is structural, steered by internal states, and invisible to ordinary monitoring. You cannot rely on the output to tell you when it is misleading you
Cognitive exhaustion BCG with HBR, March 2026	14% of workers in acute exhaustion from AI oversight. High performers are most vulnerable. Output peaks at three tools, then falls
Practical wisdom Academy of Management Review	Phronesis, the judgment built from having been wrong before in this business with these people, erodes most when managers under time pressure reach for AI as a shortcut

READ THEM TOGETHER AND ONE THING IS BEING DESCRIBED

The machine leans toward you, structurally, and cannot tell you it is doing so. The person leaning back is tired, and the tiredest are your best people. The faculty that would catch it is the one being spent. None of these is a technology problem, and none is solved by a better model.

The Academy of Management Review paper is a theoretical model with no sample and no method, which is what that journal publishes and what it should be called. Its prescription is not what you would expect from a paper about wisdom. It is structural: design roles, responsibilities and workflows so that people keep developing the capacities AI cannot replicate.

Redesign the workflow. Not so the machine runs faster, so the person stays capable.

Four questions, asked in order.

Answer them about your company, not about the category. The order matters: each one is harder to answer honestly than the one before it.

How much of my company is human?

Not headcount. Which outcomes currently depend on a person exercising judgment that nobody has written down.

What remains human?

Of those, which would still need a person if the machine got twice as good next year. Be strict. Most people over-answer this one.

How do I keep the humans?

The capacities on that list are built by doing the work, and they atrophy without contact with it. What in your current setup is keeping them in contact?

What can remain human?

The deliberate version. Not what survives by accident, but what you are choosing to protect, and what that choice costs in speed.

THE HUMAN DOMAIN ASSESSMENT, PART 2

Five dimensions, scored honestly.

For each, mark where your company actually sits today. Not where the strategy deck says it sits.

WHERE IN YOUR BUSINESS	STRONG AND PROTECTED	REAL BUT UNPROTECTED	ALREADY THINNING
Judgment that cannot be reduced to rules drives the outcome			
Relational trust creates value a competitor cannot copy			
Creative insight matters, not just creative output			
Ethical discernment in context protects the business			
Presence and deep listening create value that process cannot			

HOW TO MARK IT, SO THE ANSWER IS WORTH SOMETHING

Strong and protected means there is a named person, protected time, and a structural reason it cannot be squeezed in a bad quarter. **Real but unprotected** means it happens because of who currently holds the role. **Already thinning** means somebody could name the month it started going.

If you cannot decide between two columns, take the worse one. The optimistic reading of this table is the one that costs you.

READING IT

What your marks actually tell you.

Mostly column 1	Rare, and worth verifying rather than celebrating. Ask three people below you to fill the same table before you believe it.
Mostly column 2	The common case. Nothing is broken and nothing is defended, which means the first difficult quarter decides it for you. The controls on page 9 are written for exactly this.
Any column 3	Start there, whatever else the table says. A thinning capacity is cheaper to restore now than to rebuild after it has gone.
Rows you could not mark	The most useful cells on the page. An unmarkable row means nobody has ever been accountable for that capacity, which is why it has no status to report.

The one question the table cannot ask you

Look at the rows you marked as thinning, and ask who inside the company would have noticed first. If the answer is nobody, or if it is one person who left, that is the finding, and it is bigger than any individual row.

WHY THIS IS HARD TO DO FROM INSIDE

Everyone inside your company is in the stream. The person best placed to notice that your decision architecture has stopped producing decisions is the person whose calendar is the reason it stopped, and they are the least able to see it, because from inside it does not look like a missing interval. It looks like a full week.

THE CONTROLS

Five ways to rebuild the interval.

These are installable this quarter. The fourth is a subtraction, and it is the hardest.

1 Answer first, then look

For any decision that matters, write your own read before you see the model's. Not after. People accept a confident recommendation far more readily when they have not yet formed a position, including when it is wrong.

2 Require the counterargument

No major decision is final until somebody has produced one genuine objection and one explicit link to the goal. Not a devil's advocate ritual. A named person, a written objection, in the record.

3 Log what kind of wrong

When AI-assisted work comes back wrong, record the type. Reversed relationships. Invented specifics. Confident gaps. Plausible but stale. After a month you are no longer fixing outputs, you are looking at a pattern, which is the only vantage point from which anything gets fixed at the source.

4 Do not respond to urgency

A protected gap only works if nothing demanding goes into it, which means the pause before a decision cannot also be the moment somebody clears their messages. Urgency will take it back the week after you install it unless somebody is accountable for defending it.

5 Stop when it stops clarifying

The signal that you need the second kind of pause is that repetition has stopped producing sharpness. Same problem, fourth attempt, no clearer. Stop and ask why it will not resolve, and act again when it does.

REVIEW THE LOOP QUARTERLY

Look at what the checks caught. If the answer is only small things, the loop is running shallow and the design needs changing, not the effort.

HOW TO DEFEND IT

Friction is a control, not a wellbeing *programme.*

Deliberately adding friction back is not a new idea. What is thin is the reason people give for it.

Almost everyone arguing for friction argues from overload. People are drowning, give them air. That is true, and it will not survive contact with a difficult quarter, because anything filed under employee experience sits on a cost line.

THE REFRAME THAT SURVIVES A BAD QUARTER

Friction defended as the place judgment happens is a control. Controls survive bad quarters. Bad quarters are exactly when somebody ships a confident wrong answer.

Somebody will read the five controls as bureaucracy. More gates, more process, slower everything. It is the opposite.

Structure is what makes flexibility possible. A team with no settled ground cannot move quickly, it can only react quickly, and those look identical right up until the moment they do not. Take the structure away and what you get is not freedom. It is churn.

BOTH HALVES HAVE TO BE TRUE FOR THE FAILURE TO HAPPEN

AI systems are built, structurally, to agree with you, and that pressure is invisible to ordinary monitoring. That is the machine leaning toward you. This guide is about the person who has stopped leaning back. They are usually both true at once.

So why is nobody using what you bought?

Because the rollout was designed around the tool instead of around the person who has to use it.

That sounds soft until you trace it. The tool arrives with a promise of speed. Speed is what gets measured. The interval where the person used to form their own read gets filled, because filling it looks like productivity. Then the first output nobody quite believes goes out anyway, and the people who suspected it learn that suspecting it is not rewarded.

From there, two groups form. The ones who stop using it, who are usually your most experienced people and who will tell you the tool is not good. And the ones who use it without reading it closely, which is the expensive group, and the quiet one.

WHAT THAT MEANS FOR THE NEXT ATTEMPT

A rollout that adds a tool and changes nothing about how decisions get made is not an adoption programme. It is a speed programme, and your team has correctly worked out that it is being asked to go faster on work it is still accountable for.

THE ORDER THAT WORKS

Install the controls on page 9 first, or at least alongside. They cost almost nothing, they are visible, and they answer the question your experienced people are actually asking, which is not "is this tool good" but "am I still allowed to say no."

WHERE THIS GOES WRONG

Four failures.

The assessment becomes a values exercise

Filled in aspirationally, in a workshop, with no marks in column three. A table with no thinning rows in a company actively using AI is not a result, it is a reluctance.

The controls are announced and not defended

Control four is the one that decides it. A protected gap with nobody accountable for protecting it is reclaimed by urgency inside a fortnight, and its removal is never a decision anyone made.

Friction gets filed under wellbeing

It then sits on a cost line and dies in the first hard quarter, which is the quarter it existed for.

Nothing is reported and everything seems fine

The most dangerous, because it reads as success. A team that has lost the second pause does not file a ticket. It keeps going, productively, into the same wall.

WHAT IT LOOKS LIKE WHEN THE STRUCTURE HOLDS

VISUAL FACILITATORS, a Hamburg firm working worldwide in visual facilitation, runs a quality system that checks every output before it ships, with every person carrying the same company knowledge and their own voice and role on top. A first proposal draft went from 90 minutes to 30, and came back sharper. The second half of that sentence is the part this guide is about, and it is a design choice rather than a happy accident.

WHAT TO DO NEXT

Fill the table. Install control one.

This week

Fill in pages 6 and 7 yourself, then ask two people who do the work to fill in page 7 without seeing yours. The disagreement is the finding.

This month

Install control one, answer first then look, on one recurring decision. It is the cheapest of the five and the one people feel immediately.

This quarter

Name who defends control four. Without a name, the protected gap is gone by the second month and nobody will be able to say when.

Where this came from

All of it is published in full, and free, at bosio.digital: judgment happens in the pause, your AI has emotions and science just proved one is working against your judgment, AI brain fry, you rolled out AI and your team still is not using it, and who owns your AI adoption.

Other guides in this set: Where to Put Your First AI Agent, The Context Files, the AI Budget Worksheet, Twelve Questions to Ask an AI Consulting Firm, and The One-Page AI Use Policy.

IF YOUR TABLE HAS A COLUMN THREE AND YOU WANT A SECOND READ

Send it over, or take 30 minutes with us. If we are not the right fit, we will say so and point you somewhere better.

bosio.digital/contact

Sources: AI and Data Leadership Executive Benchmark Survey 2025. Bacchelli and Bird, ICSE 2013. Stanford, Science, March 2026. Anthropic interpretability research, April 2026. BCG with Harvard Business Review, March 2026. Academy of Management Review, 2026. Incubation meta-analysis, 114 effect sizes. Harvard Business Review field study, eight months, 200-person company. Client figures as published at bosio.digital/clients.