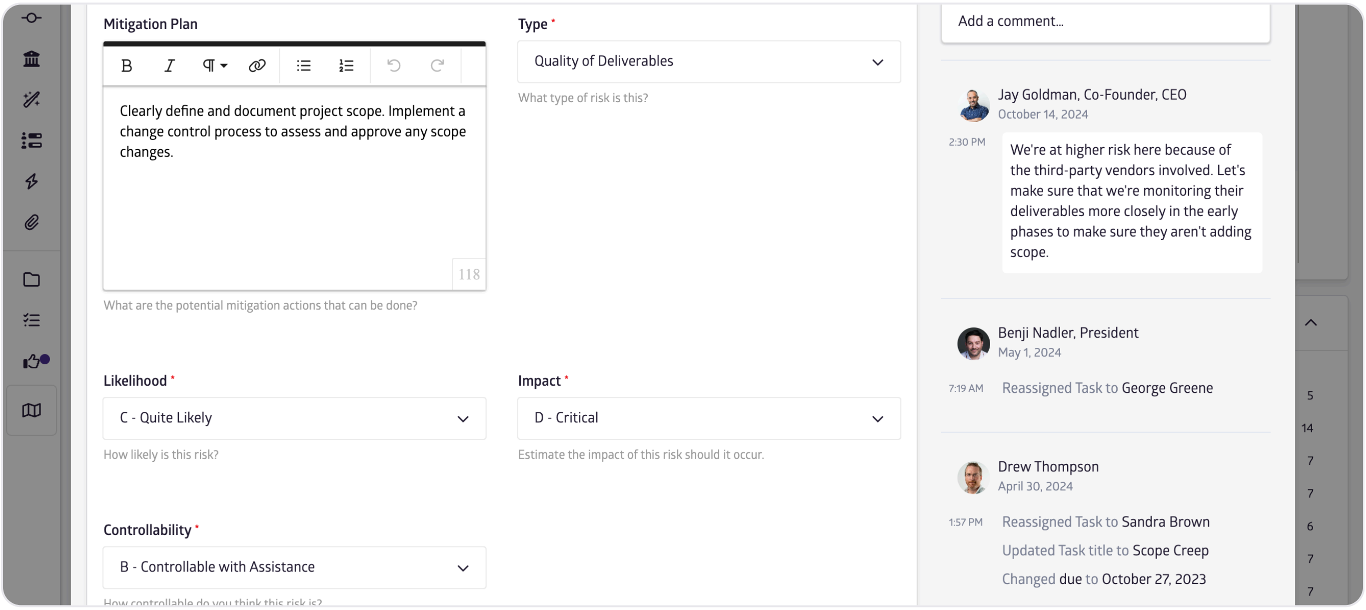
THOUGHT LEADERSHIP

The Most Durable Problem in Management

For two decades, companies have captured only about two-thirds of the value their strategies promise. Every method we have thrown at that gap helped at the edges, and none of it closed the gap. A layer of AI agents may be the first thing that does, and it changes what the strategy execution office is for.

Jay Goldman July 27, 2026



Here is what I keep seeing. I spend most of my time with the people who run transformation inside large companies, and lately the same conversation keeps happening. They can tell you, in detail, what a strategy is supposed to deliver. Ask what actually landed, and the number is always smaller, and nobody is quite sure where the rest went.

That distance, between what a strategy promises and what an organization gets, is the most durable problem in management. It has outlasted every remedy prescribed for it. The program office. The balanced scorecard. The transformation methodology. The quarterly objective. Each helped a little. None of them closed the gap. Companies still capture about two-thirds of the value their strategies set out to produce, roughly what they captured when the figure was first measured twenty years ago.

So a little skepticism is healthy when a new technology shows up claiming to fix what nothing before it could. That is more or less the claim now being made for AI agents, and most of it is overstated. But a narrower version of the claim is true. And the narrower version is enough to change what the strategy execution office is for.

WHY IT NEVER CLOSED

Why the gap survived

Start with why the gap has been so stubborn. Underneath, the failures of execution are failures of coordination and information. A priority that was never honestly ranked. A benefit no one measured after the business case was signed. A program slipping quietly for six weeks before the review noticed. A dependency that fell into the space between two teams. None of this is intellectually hard. It is merely relentless, the kind of continuous, unglamorous attention no leadership team has the hours to sustain, and that the program office, invented to supply it, mostly converts into status reports.

The work that would actually close the gap has always existed. There has simply never been enough of it, because it is expensive human labor and there are only so many analysts, so many hours, so many reviews in a quarter. That scarcity, not a shortage of strategy or talent, is what the missing third is mostly made of.

I have sat in enough of these reviews to know the rhythm by heart. The first half goes to agreeing on what is even true. The decisions that would actually move the number get whatever time is left, which is usually not much.

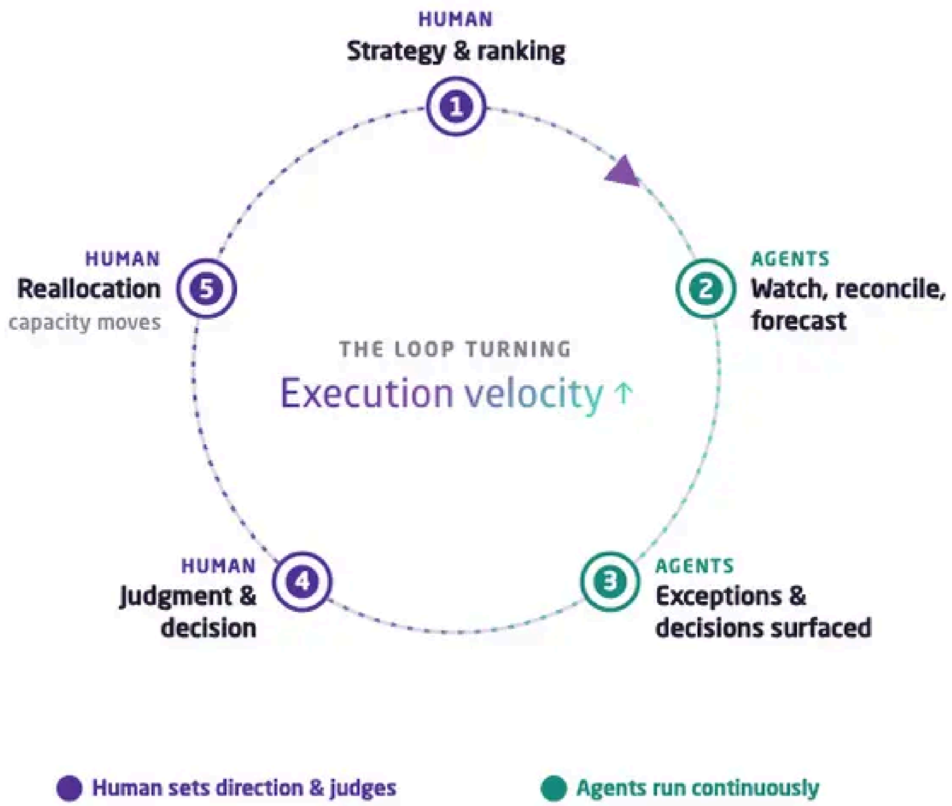
THE SHIFT

What an agentic execution office is

An agentic execution office is what emerges when that work stops being scarce. Set aside the science-fiction version, the one where an autonomous system runs the company while the executives play golf. The useful version is more modest and much nearer to hand. It is a layer of software agents that continuously performs the gathering, reconciling, monitoring, and prompting the execution office is meant to do and has never had the capacity to do well.

EXHIBIT 1

The agentic operating loop



Read it this way. The agents read the systems where work actually happens, the financials, the delivery tools, the tickets, the documents, the message threads, reason over what they find, and either surface it to a person or, within limits agreed in advance, act on it. The office keeps the part that was always its own and always crowded out by administration: judgment. What changes is not the intelligence in the room, but the amount of attention available to spend, which rises from whatever a stretched team could manage to something close to continuous. **Turquoise nodes run continuously. Purple nodes stay human.**

The reason this matters is not that judgment gets automated. Judgment was always the human part, and it stays the human part. It is that the continuous attention around the judgment, the part that was always too expensive to buy in the quantity the job actually needed, stops being scarce. That is the whole shift, and it is worth being precise about what it costs today.

I will be honest about something we learned building toward this. I assumed the hard part would be the technology. It was not. The models are ready, and more capable than most of the people using them realize. The hard part is everything upstream of them, which is where the rest of this piece goes.

THE COST OF THE GAP

What the missing third is made of

The price of running the company's biggest bets on attention no one could afford is not hidden, though we rarely say it plainly.

EXHIBIT 2

The gap, four ways

63%	**1/6**	**72%**	**~10¢**
of a strategy's value is realized on average. About a third leaks in execution, and the number has barely moved in twenty years.	major IT programs becomes a "black swan," an overrun averaging around 200 percent, in a database of 16,000+ projects.	of PMOs still spend half a day or more every month just assembling reports by hand.	of every dollar is wasted on poor project delivery, roughly one dollar in ten.
MANKINS & STEELE, HBR	FLYVBJERG, OXFORD	WELLINGTONE, 2026	PROJECT MGMT INSTITUTE

None of this is a strategy problem. The strategies are usually sound. It is an execution problem, and underneath that, an attention problem. Value leaks at the handoffs, in the space between what leadership committed to and what the organization can actually see, week to week, without someone staying up to watch it.

THE METHOD

The four gaps, worked continuously

The value of continuous attention gets concrete when you point it at the four places value leaks: clarity, alignment, cadence, and insight. Closing them never took more intelligence in the room. It took more attention than any team could afford, applied without pause. That is exactly what a layer of agents can now supply.

EXHIBIT 3

The four gaps, worked continuously

01 Clarity

An agent keeps the ranked priorities and their intended benefits legible, and flags the new request that maps to none of them. It cannot decide what matters. It can stop the portfolio sprawling back to forty things without anyone noticing.

02 Alignment

It watches whether funding, capacity, and ownership actually sit behind the stated priorities, and surfaces the program everyone calls strategic that is quietly staffed by two contractors. Those discoveries stop happening by accident, and late.

03 Cadence

It assembles what moved, what is blocked, and the decision each owner needs, so the review stops being a status ritual and becomes a forum for decisions. A slipping dependency escalates the day it slips, not at the next steering committee.

04 Insight

It measures benefits against the business case continuously, from the real operational and financial data rather than a self-reported register, and flags the program falling below plan the week it starts to, not at the closure review.

The pattern underneath all four is the same. The work that closes the gap was never intellectually hard, only relentless, the continuous attention no team ever had the hours to sustain. That is the scarcity agents lift, and continuous measurement is what finally makes continuous reallocation possible.

WHAT IT LOOKS LIKE

A week in the office

None of this is hypothetical in the way it would have been three years ago. Picture the office on an ordinary Tuesday.

Overnight, a benefits agent reconciled the portfolio's claimed benefits against the source data and found one program running eight points below its business case. It drafted a one-page exception with the likely cause and three options. A dependency agent noticed that a data-migration workstream is a week from blocking two others, and already routed the question to the person who can re-sequence it. A clarity agent flagged a new request from a business unit that arrives with a healthy budget and no connection to any of the year's priorities.

The transformation director opens a single view that already knows what changed and why. The morning goes to three genuine decisions, whether to intervene on the slipping program, how to clear the dependency, whether to admit the unranked request, instead of the three days that assembling the same picture would once have taken. The executive review later that week runs for forty minutes and produces four decisions, because the status was settled before anyone sat down.

THE NUMBER THAT MATTERS

Execution velocity

There is a single measure underneath all of this, and it is the one I think this work should be judged on. Call it execution velocity: how fast, and how completely, an organization turns strategy into realized value. Not activity. Not delivery throughput. Realized value, and the time it took to get there. Every one of the four gaps an agent helps close is in service of the same thing, raising that number.

I have started asking almost every leader I meet a version of the same question: what is your execution velocity right now, and is it going up? Almost no one can answer. The best ones go quiet, because they know they should be able to, and that not knowing is itself the problem.

THE NORTH-STAR METRIC

$$\text{Execution Velocity} = \frac{\text{priority initiatives} \times \text{realization rate} \times \text{average value per initiative}}{\text{time to benefit}}$$

More of the right work, more of its value actually realized, delivered sooner. Most companies have never measured it. It is the closest thing this work has to a north star, and it is the number a continuously attended portfolio is built to move.

THE ROAD THERE

Four stages, and the one that matters

The shift will not arrive whole. Most organizations sit at the first of four stages, and the distance between them is mostly a question of how much the office is willing to trust something other than a person to notice.

EXHIBIT 4

The road there, in four stages

01 · ASSISTIVE	Agents draft the status and summarize the register. Copilots for the analysts. Useful, and a long way from the prize.
02 · CONTINUOUS	Agents monitor benefits and dependencies unprompted and raise exceptions between reviews. This is the jump that matters.
03 · ORCHESTRATING	Agents run the cadence end to end, composing the review and routing each decision to the person who must take it.
04 · BOUNDED AUTONOMY	Within limits agreed in advance, agents re-sequence work and reallocate capacity on their own, bringing only the exceptions to a human.

Most offices sit at stage one. The leap worth making is from one to two, from a cleverer way of producing reports to a system that watches the portfolio when no one is looking. Most of the value lives there, and so does most of the difficulty, because it is the point at which the office has to trust something other than a person to notice.

WHAT STAYS YOURS

Rank first, then automate

The temptation with a capable tool is to point it at everything, and here that instinct is genuinely dangerous. An agentic office built on an unranked portfolio and a set of fictional business cases will not fix them. It will leak value faster and with more conviction, and produce beautifully evidenced reports on the way down. Someone still has to decide what matters, what it is worth, and what to stop. An agent can enforce a ranking. It cannot invent one. Rank first, then automate. Automation applied before clarity just industrializes the confusion.

EXHIBIT 5

Where the line sits

THE AGENTS TAKE

Continuous, and answerable to a human

- Gathering status and reconciling the numbers
- Watching benefits against the business case
- Catching a slipping dependency the day it slips
- Flagging the request that maps to no priority
- Composing the review before anyone sits down

YOU KEEP

The consequential, human calls

- Deciding what matters, what it is worth, what to stop
- Killing a program with a powerful champion
- Moving scarce specialists against the politics
- Holding an executive to a benefit they missed
- Governing the agents, and answering for them

The office picks up a duty it did not have before: governing the agents themselves. What they may do unattended, what they must bring to a person, and who answers when one of them is confidently wrong. An office that runs on agents is accountable for them, and that responsibility does not automate.

THE COMPOUNDING

The agents are the easy part

The strategy execution gap endured because closing it demanded an amount of continuous attention no organization could afford to buy. That constraint is the one now lifting. An office that takes the opening turns itself from a function that reports on the past into an engine that compounds realized value, and the compounding is the whole point. When a competitor's strategy lands two-thirds of what it promised and yours lands most of it, the gap between the two companies does not stay modest for long.

The technology to do this is arriving whether or not any given office is ready for it. The work now is to become ready, to rank the portfolio honestly, wire the benefits to real data, and build the judgment and the guardrails a system this capable will demand of the people who run it. Those are the hard parts, and they are entirely human. The agents are the easy part.

WHERE TO START

Three moves, not a framework

If you want to start, and I think the case for it is strong, you do not need to reinvent the whole office first. You need three moves.

Rank honestly. Force the portfolio down to the few priorities that actually carry the strategy, and stop or park the rest. This is the human act the whole thing rests on. An agent can hold the line once you have drawn it. It cannot draw it for you.

Wire one benefit to real data. Pick a single flagship program and connect its business case to the systems where the work and the money actually move, so its health is measured continuously instead of self-reported once a quarter.

Run the loop on that one program. Put agents on the continuous watching for it, keep every consequential call with your people, and let what you learn there set the pattern for the next program, and the one after that.

That is the whole move from stage one to stage two, and it is less than you fear. It is also the only version of this that compounds: the standard you set on the first program is what makes the next one faster, and the one after that.

WHERE I AM GOING WITH THIS

An invitation

This is the problem my team spends its days on. We are getting close to launching the next version of Conductor, and everything in this piece is built into it. The closer we get, the clearer one thing becomes: the opinion that matters most to me now is not another analyst's. It is yours, from the people who actually run this work every day.

So if you are curious to see how we are building it, reach out and I will walk you through it myself. I would rather put it in front of a handful of transformation leaders whose judgment I trust, while there is still time to get it right, than unveil it finished and hope we did. You understand this work better than any roadmap ever will, and I would like your fingerprints on it before it ships.

Either way, stay with me here. This newsletter is where I will keep sharing the thinking as we build, one piece at a time. And if you want the closer look, or you simply want to tell me where I have this wrong, reply below or send me a message. I will follow up myself, and I read every one.

— Jay

READ ONLINE

Read this article on our website

getconductor.com/resources/the-most-durable-problem-in-management

GET STARTED

Start a trial or talk to sales

Try Conductor yourself or walk through an example with our team.

Start a trial →

Talk to Sales

Scan to start