UPwind

A WORKING POSITION · JULY 2026

AI governance as a portfolio, not an approvals gate

The tool you approved doesn't stay the tool you approved.

Tim Jones · UPwind

Where the standard playbook breaks

Enterprise IT governance has a standard shape: assess the risk, approve the tool, record the decision, move on. It works because the thing being governed holds still. Infrastructure is stable, versions are knowable, and an approval written in March still describes reality in November.

AI breaks that in one specific way. The tool you approved does not stay the tool you approved. Vendors swap the underlying models, capabilities appear between releases, pricing gets restructured, and none of it asks permission or sends a memo. The approval stands while the thing it described moves. A register of approved tools is accurate the day it is written, quietly wrong six months later, and issuing assurance the whole time.

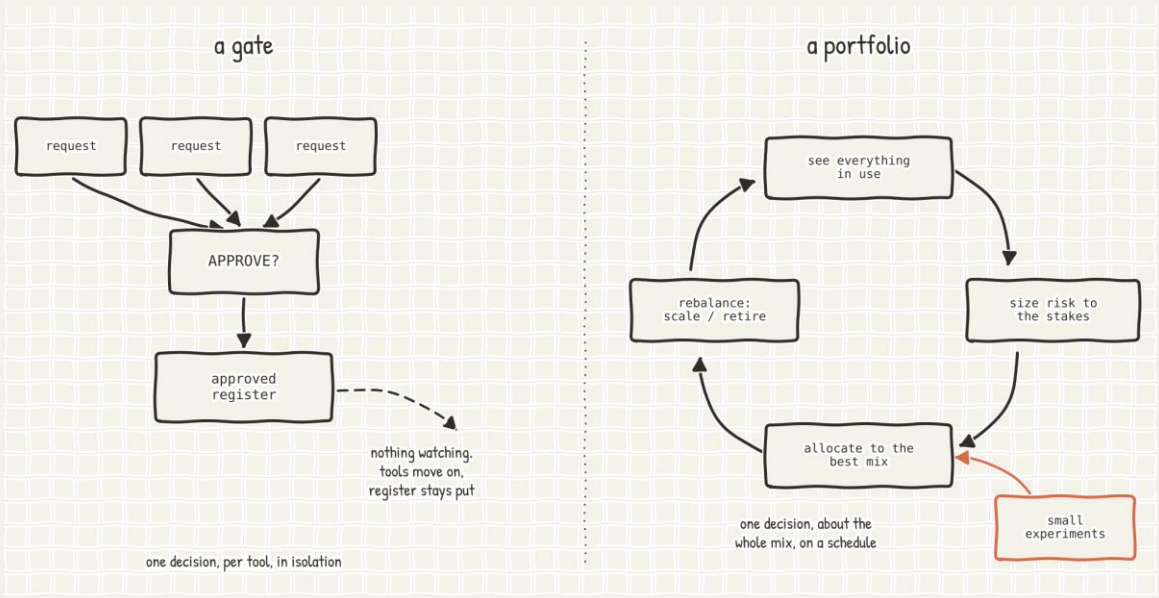
There is a second, quieter failure. An approvals gate does not stop unsanctioned use; it hides it. Wherever there are staff, browsers and deadlines, AI tools are already in use, approved or not. So the practical choice is not between AI with governance and no AI. It is between use you can see and use you cannot.

The third failure costs the most and gets noticed last. A gate prices every request as a finished business case: a defined use, a predicted benefit, approve or decline. AI value rarely arrives in that shape.

Even the defined use is shakier than it looks, because the return lives in the pairing of tool and person rather than in the tool. The same licence is a different investment on different desks: in one team it pays for itself inside a week, and two desks over it is a subscription nobody opens. No form captures that. The value that does arrive turns up in use, usually somewhere nobody predicted, and it usually starts as a small unofficial experiment that could never have justified itself on paper. A gate tuned to certainty screens those out, which means it filters out the discovery along with the risk.

What a portfolio actually means

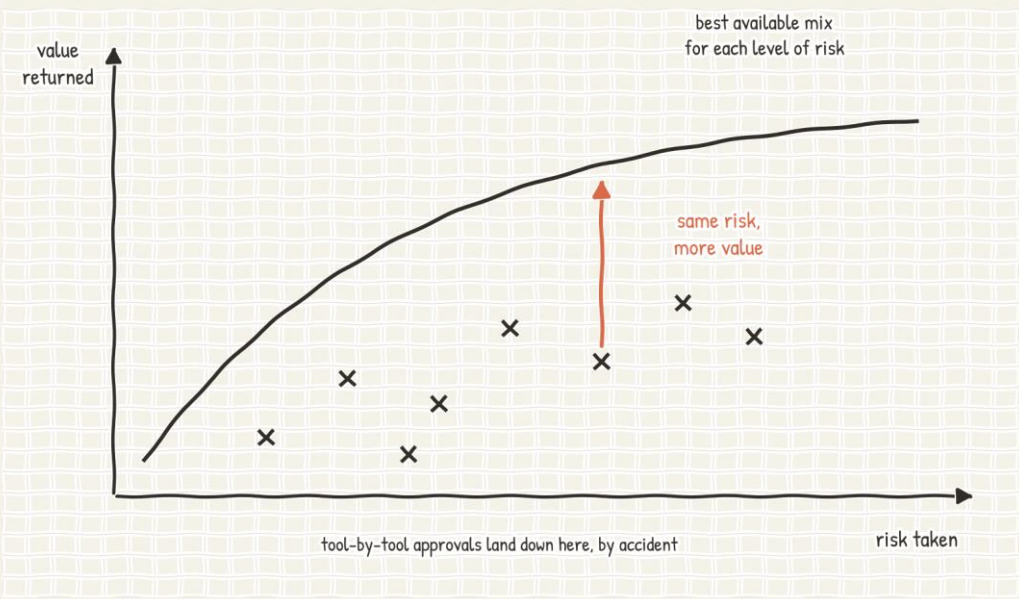
A gate with a review calendar attached is still a gate: the unit of decision is one tool, judged on its own merits, yes or no. A portfolio changes the unit of decision. The menu of AI applications a business could run is now effectively unlimited. The scarce things are the risk you are willing to carry and the money and attention you can commit. So the governing question stops being whether this tool is acceptable, asked one request at a time, and becomes: for the risk and capital we are prepared to commit, what mix of positions returns the most?



Once the unit of decision is the mix, things become visible that no per-tool approval can see.

- Overlap: three teams separately paying for four tools that do the same job.
- Concentration: everything riding on one vendor.
- Gaps: no exploratory positions anywhere.
- Sizing: large commitments to proven uses and small stakes on unproven ones, instead of one uniform bar every request must clear.

An investment portfolio manager already has a picture for this. For any level of risk an investor is willing to hold, there is a best available mix of holdings, and every other mix returns less for the same risk. Swap the investor for a business and the picture holds. Tool-by-tool approvals land somewhere under that line, by accident rather than by anyone's failure: risk gets taken wherever it happens to be requested, which is rarely where it returns the most.



The gap between those two positions is the whole argument. The business is carrying the risk either way; managed one approval at a time, the same risk simply buys less value than it could. Closing the gap means getting properly paid for risk the business is already holding.

One caution before anyone opens a spreadsheet: the frontier is a mental model, not a chart to compute. Professional fund managers do not plot one curve across everything they hold either. It is a way of seeing the space, and it earns its keep through the two habits it forces: pricing every position in the same currency, and sizing bets to their risk.

The obvious objection is that the positions look nothing alike. How do you weigh a vendor compliance agent against a small sales automation? The same way an investor weighs a bond against a stock: by what each does to the whole mix. The trick is one currency. Value is dollars a year, and loss avoided counts as value, which is how the compliance work gets on the board at all. Risk is expected dollars of loss a year: how often something goes wrong, times how much it costs when it does. Cyber security teams have priced unlike risks this way for two decades under a method called FAIR, built precisely so that a data breach and a payroll error could sit in the

same conversation. A simplified *Loss vs Frequency vs How-Long-Have-We-Been-Running-Our-Own-Trial* calculation works fine as well.

A position that cuts expected loss also does a second job: it frees up risk budget to spend on more aggressive positions elsewhere. Nobody asks whether the smoke alarm out-earns the espresso machine. The smoke alarm is what lets you run the espresso machine in a wooden building.

A real portfolio also holds a sleeve of small speculative positions, deliberately, because that is where discovery happens. In AI terms these are the pilots: contained, visible, cheap experiments with a route to being scaled when one proves out. That sleeve is the research arm of the portfolio, and it is the part a gate structurally cannot provide. A gate can only decline uncertainty. A portfolio prices it, sizes it small, and watches for the one that pays.

The working rhythm

None of this needs a big apparatus. What it needs is a handful of habits, most of which any business already runs somewhere else under other names.

Visibility first. You cannot allocate what you cannot see. One discovery pass builds the book: every tool actually in use, approved or not, what it does, what data goes into it, what it costs, and what it duplicates.

Then the book stays honest: unapproved use keeps getting recorded rather than punished, because the moment a log entry carries a penalty people stop making them, and the exposure goes back underground. The gap between the book and the approved register is the current risk position. Until the book exists, the exposure is simply unmeasured.

A risk budget, not a uniform bar. Decide explicitly how much risk the business is prepared to hold, and where. The arithmetic is the one every risk framework already uses: how often a thing goes wrong, times how much it costs when it does.

The step that usually gets skipped is the baseline. Every AI position replaces or supplements a process run by people, and that process already fails at some frequency and some severity; nobody has priced it, because it was never on a register. That baseline is what the agent has to beat, and it is rarely zero errors. Customer-facing, regulated and decision-making uses get the heavy treatment. Pilots get light rules and hard walls.

Rebalancing, on a trigger, with an owner. A review calendar on its own recreates the gate with a date attached. Rebalancing should move when the world moves, and the triggers are knowable: a major model release, a big capability rollout on a platform already in the book, a competitor shipping something, with a regular cadence underneath as the backstop so nothing drifts for a year.

The questions are portfolio questions rather than per-tool ones:

- What is earning its keep.
- What overlaps.
- Which pilot proved out and deserves scaling.
- Which holding gets retired because something better has been released.

An hour usually covers it. The owner matters more than the calendar.

Lanes people will use. A rule people route around converts visible experimentation into invisible exposure, so the sanctioned path has to be quicker than the workaround. In practice this is a pilot lane: a deliberate, fast way to run a small experiment inside hard walls, with the data it can touch and the money it can spend capped tightly enough that the honest answer to the worst-case question is: not much. Anywhere else in the business these would just be called pilots and trials.

The portfolio treatment adds the two pieces usually missing, a size limit on the way in and a known route up: an experiment that proves its value gets promoted into a funded position with proper controls, and one that does not gets closed without ceremony. Add an amnesty for what is already in use and a short approved list for the common cases, and the workaround stops being the fast lane. The promotion route is how value discovered at the grassroots actually reaches the portfolio.

Where it starts

The starting state is small: the discovery pass to build the book, the pilot lane with its walls, and the first rebalance in the diary with an owner and a trigger list.

From there the discipline is position sizing, which is all a portfolio really asks: take bets in proportion to their risk:

- How often it can go wrong
- How much it costs when it does
- How much of the estimate is still guesswork.

A new agent's numbers are estimates, so they carry a discount until real runs replace the guessing; an agent that has run cleanly a hundred thousand times has earned a trust that an agent run once has not, whatever its theoretical risk says. Insurers have priced new drivers on exactly this logic for a century. Trust is lent at the start and earned after.

The expensive failure here is not the tool that goes wrong. It is the hundred small experiments that never happen because the gate made them not worth the bother.