

The Operating Model Gap: Why Tool-Led AI Fails

Whitepaper

Bart de Graaff · BdG Advisory · October 2026

Tools can be bought in a quarter; ownership, incentives and decision control cannot. That is why AI usage spreads while enterprise results stay flat. This paper names the three places the gap opens, why each is rational for the people inside it, and the order in which leaders close it.

The tool is not the transformation

Most organisations are buying AI. Very few are changing how they decide.

That gap is where the money goes.

The structural reality is simple to state and hard to act on. AI enters through tools: a licence, a platform, a copilot, a pilot. Value only appears when the operating model around the tool changes: who owns the work end to end, what people are paid to do, and who is accountable when a system makes the call. Tools can be bought in a quarter. Structure, incentives and control cannot. So organisations buy what they can buy, measure what they bought, and then wonder why execution does not move.

The evidence is consistent. McKinsey's March 2025 global survey found that more than 80 percent of respondents see no tangible impact on enterprise-level EBIT from generative AI. The same survey tested 25 organisational attributes, and the one with the biggest effect on EBIT impact was the redesign of workflows. Only 21 percent of respondents whose organisations use generative AI said they had fundamentally redesigned even some of them. A widely cited 2025 MIT NANDA study put it more bluntly: about 5 percent of AI pilots achieve rapid revenue acceleration, and the rest stall with little measurable effect on the P&L. The authors did not blame the models. They blamed a learning gap: tools that do not adapt to the workflows they are dropped into.

None of this is a technology problem. It is an operating model problem that technology makes visible.

This paper is the entry point to the series. It names the three places the gap opens (structure, incentives and control), explains why each is rational for the people inside it, takes on the strongest argument for letting tools lead, and sets out the order in which leaders close the gap.

Why tools come first

Tool-led adoption is not stupidity. It is the rational output of how large organisations allocate money and credit.

A tool is visible. It can be demonstrated in a steering committee. It fits a procurement category, a budget line and a vendor's sales motion. It produces a dashboard of licences issued and active users that a board can read. It lets every function say it has an AI programme without asking any function to give anything up.

The operating model is the opposite. It is invisible until it fails. It has no vendor. Changing it means moving budget, authority and headcount between people who all report to the same executive committee. Nobody gets promoted for a decision-rights matrix.

So each decision-maker does the sensible thing from where they sit. The CIO buys a platform, because platforms are what IT is accountable for. Business units buy copilots, because they can expense them. Procurement negotiates the unit price, because that is how procurement is measured. Employees do not wait. Microsoft and LinkedIn's 2024 Work Trend Index found that 75 percent of knowledge workers already use AI at work and 78 percent of those users bring their own tools. In the same survey, 60 percent of leaders said their company lacks a vision and plan to implement it.

Read those two numbers together. Adoption is running ahead of design. The organisation has the technology and no structure to direct it.

We have seen this pattern before. Cloud migration was often "break the process, move it to the cloud". Applications moved, the operating model stayed, and the savings were smaller than the business case. AI is repeating the move with higher stakes: break the process, add AI.

Consider a large hospital network that licenses an AI assistant for ten thousand staff. Within six months the adoption dashboard is green. Administrators draft letters faster. Clinicians summarise referrals in minutes. Yet waiting times, length of stay and cost per admission are where they were. Nothing failed. The tool did exactly what it was bought to do. The organisation simply had nowhere to put the time it saved, and nobody whose job it was to find somewhere.

That is the operating model gap in its purest form. The technology works. The structure absorbs it without changing.

There is a second cost to tool-led adoption, and it compounds. When the tool arrives before the structure, the tool's design fills the vacuum. The vendor's workflow becomes the default workflow. The vendor's data model decides which information is captured and where it lives. The vendor's roadmap decides what the organisation can do next year. None of this is sinister. A product has to assume an operating model in order to work, and if the buyer has not defined one, the product's assumptions win.

Each technology wave has redistributed control in this way. The internet rewarded aggregation. Cloud concentrated infrastructure control and introduced consumption economics. AI is shifting value towards whoever controls models, data and distribution, and that structure is still forming. An organisation that adopts tool by tool is making positioning decisions in that ecosystem without noticing it is making them. The decision that matters is less which technology to adopt than where to stand in the system that technology creates, and how to align the operating model to that position.

Structure: the tool lands in work nobody owns end to end

Value in AI sits across handoffs. Structure in most enterprises sits inside functions.

A customer request, a credit decision or a supply-chain exception travels through sales, operations, risk, finance and IT. Each function owns its step. Nobody owns the flow. When AI arrives, it gets inserted where someone has authority to insert it, which is inside one step. The step gets faster. The flow does not.

This is the mechanism behind McKinsey's workflow finding. Redesigning a workflow requires authority over all of it. Most organisations have not assigned that authority to anyone below the chief executive. So AI is layered onto the existing process, and the existing process sets the speed.

Take a regional bank's business lending team. An AI system now drafts credit memos in twenty minutes instead of two days. The credit committee still meets on Thursdays. Risk still wants the memo in the old template. Legal still reviews covenants in sequence, not in parallel. Time-to-decision for the customer barely moves. The local improvement is real and invisible, because the constraint was never the memo. It was the structure of approvals around it.

Data follows the same structure. AI is most useful when it can draw on information that crosses functions: the customer's history in sales, their payment behaviour in finance, their complaints in service, their usage in operations. Each of those data sets has an owner, a definition and an access policy designed for its own function. Nobody owns the joined-up view. So the AI team spends its first year negotiating access, reconciling definitions and building extracts, and the model ends up trained on whatever one function was willing to share. Data reflects how decisions are made. If ownership is unclear, data fragments, and AI amplifies the fragmentation rather than fixing it.

Worse, local speed creates new queues. When one step accelerates and the next does not, work piles up at the boundary. People respond the way people always do: they build workarounds, side channels and exceptions. The organisation becomes busier and no faster.

This pattern is not unique to AI. It recurs across operating environments. In venture-backed companies, commercial pressure turns structured product roadmaps into deal-driven priorities. Revenue grows, but delivery complexity grows faster. In large enterprises, governance introduced to manage risk expands until it slows every decision and dilutes

accountability. Capital gets deployed ahead of delivery capability, creating commitments the organisation cannot meet.

The result is predictable.

Technology capability increases. Execution reliability decreases.

AI intensifies this. By raising the speed of individual steps and removing manual effort, it strips out the slack that used to hide structural problems. Manual processes were slow, and slowness gave people time to compensate for unclear ownership. Remove the slowness and the questions that could be deferred become immediate: who owns this decision, in what sequence does work happen, do our commitments match our capacity, and who controls the data.

These questions are not new. AI makes them impossible to ignore.

Incentives: people are paid to run the old model

Structure decides who *can* change the work. Incentives decide who *wants* to.

Every operating model encodes a set of rewards. Managers are measured on budget delivered, headcount managed and utilisation. Status follows the size of the team. Suppliers are paid for effort, licences or seats. Sales organisations are paid on this quarter's revenue from existing products. None of these measures was designed with AI in mind, and almost all of them reward the volume of work rather than the outcome of it.

Now introduce a system that removes a third of the effort in a function. For the function head, that is a smaller budget next year and a smaller team. For a delivery partner priced on time and materials, it is less billable work. For a sales team, it is a new product that takes longer to close than the old one. Each of them responds rationally. They support the pilot, praise the technology and protect the economics it threatens.

The budget cycle reinforces it. In most enterprises, a saving identified this year is removed from next year's budget. Finance captures the efficiency as a cut, and the manager who found it ends the year with less. Managers learn quickly. They stop surfacing capacity and start absorbing it: the team takes on more projects, the reports get longer, the backlog shrinks a little. The AI saving is real, but it never becomes visible enough to be taken. A structure that punishes the people who reveal efficiency will not get much of it revealed.

This is why resistance to AI rarely looks like resistance. Nobody blocks the programme. They slow the scaling, expand the governance, ask for one more proof of concept and redeploy the saved capacity into work that preserves the team. The pilot succeeds. The P&L does not change.

The same dynamic shows up in partnerships. Large sales organisations optimise for existing revenue. A strategic alliance with a technology provider creates brand association, but brand association does not create priority, and access does not create demand. Partnership only

works when senior sponsorship is active, targets are mandatory and individual upside is tied to execution. AI programmes run through partners fail for the same reason.

The incentive problem does not stop at middle management. It runs to the top. Boards ask what the AI strategy is, and investors listen for it on earnings calls. An executive can answer that question with a platform announcement, a partnership and a count of use cases in a single quarter. Answering it with a redesign of decision rights takes two years and produces nothing to announce until it works. So executive attention is rewarded for the visible part of the programme and rarely for the structural part. The organisation learns what leadership actually values from what leadership talks about.

Leaders who capture value act on this directly. BCG's October 2024 survey of 1,000 executives found that 74 percent of companies had yet to show tangible value from AI. The leaders followed a different allocation: about 70 percent of their resources on people and processes, 20 percent on technology and data, and only 10 percent on algorithms. That ratio is an incentive statement. It says the hard part is changing what people are rewarded for, not choosing the model.

Consider a contact-centre operation inside a utility. The AI pilot cuts average handling time by a quarter. The centre's manager is measured on service levels and cost per contact, and her standing in the organisation is tied to running the largest operational team in the business. The pilot is extended, celebrated and never scaled beyond two queues. Nobody acted in bad faith. The structure paid her to keep the team.

Incentives alone do not fix this. Paying people a bonus to adopt AI while leaving their budget, headcount and status measures unchanged buys the appearance of change. Carrot without structure is bribery. The measures themselves have to move.

Control: when the system decides, someone must own the decision

The third gap is the most dangerous, because it stays invisible until something goes wrong.

Enterprise governance was built to approve software. A change board reviews a release, signs it off and moves on. The logic inside the system was written by people and can be traced to them. AI does not fit that model. It produces a continuous stream of decisions, many of them small, some of them consequential, none of them individually approved. Governance designed to approve a tool has no mechanism to own a stream of decisions.

The public record already shows what happens next.

In February 2024, a Canadian tribunal ruled on *Moffatt v. Air Canada*. The airline's website chatbot had told a grieving customer he could claim a bereavement fare retroactively, which was not the airline's policy. Air Canada argued that the chatbot was a separate legal entity responsible for its own actions. The tribunal rejected this and held the airline responsible for all the information on its website, whether it came from a static page or a chatbot. The amount at stake was small. The structural point is not. An organisation cannot deploy a

decision-making system and then disown its decisions.

Zillow offers a larger version. Its home-buying business, Zillow Offers, bought houses directly at scale, relying on its own forecasts of what they would sell for. In November 2021 the board decided to wind the business down, citing in its SEC filing home pricing unpredictability, capacity constraints and other operational challenges. The company recorded a \$304.4 million inventory write-down for the third quarter and expected to cut about 25 percent of its workforce. Forecasting house prices is hard, and no model would have been perfect. The structural lesson lies elsewhere: a forecast was turned into balance-sheet commitments at the speed of an operating programme. The design question that matters is who had the authority, and the trigger, to slow the buying when the model's error exceeded the margin.

Australia has the starkest example, and it was not even machine learning. The Commonwealth's automated debt-raising scheme, known as Robodebt, was examined by a Royal Commission that reported on 7 July 2023 with 57 recommendations. Among them: a consistent legal framework for automation in government services; a clear path for people affected by automated decisions to seek review; business rules and algorithms made available for independent expert scrutiny; and a body with power to monitor and audit automated decision-making. Every one of those recommendations is a control structure that should have existed before the system ran.

Read the three cases together and the pattern is clear. The technology did roughly what it was built to do. What was missing was a named owner of the decisions, a defined boundary on what the system could commit the organisation to, a mechanism to stop it, and a trail that let someone outside the team check its logic.

Leadership attention matters here more than in any other part of the gap. McKinsey found that a CEO's oversight of AI governance was one of the elements most correlated with higher bottom-line impact from generative AI, and at larger companies the element with the most impact. Control is not a compliance afterthought. It is the condition for scaling at all. The cognitive governance perspective takes this further, into how AI decision-making itself should be contained.

The strongest objection: let adoption lead

The best argument against everything above goes like this.

Redesigning the operating model before deploying AI is slow, expensive and probably wrong. The technology changes every few months. Any end-state design will be obsolete before it is implemented. Meanwhile, employees are already using the tools and finding value on their own. The personal computer, email and the spreadsheet all spread without anyone redesigning the operating model first. Let adoption run, watch where value emerges, and formalise it afterwards. The MIT study even found that buying specialised tools from vendors succeeded far more often than internal builds. Buy good tools, get out of the way.

Much of this is right. Designing the full end state in advance is a mistake. Bottom-up use does surface real opportunities. Buying usually beats building for anything that is not a source of advantage.

But the conclusion does not follow, for three reasons.

First, individual productivity does not become organisational value on its own. When a tool saves a person an hour, the hour is absorbed by the existing structure: more email, more meetings, more polishing. It reaches the P&L only if someone with authority redesigns the work, the capacity plan or the budget to capture it. That is why adoption can reach three-quarters of the workforce while more than 80 percent of organisations see no enterprise-level EBIT impact.

Second, unmanaged adoption accumulates unowned risk. When staff bring their own tools, customer data, decisions and commitments flow through systems nobody has assessed. Air Canada shows where that ends: the organisation owns the outcome whether or not it designed the control.

Third, the historical analogy cuts the other way. The general-purpose technologies of the past did eventually transform organisations, but only after long periods of reorganisation, and the gains arrived when the structure changed, not when the tool arrived. The companion paper, *The Last 20%*, examines that absorption problem directly: what happens when AI works and the organisation still has to be rebuilt around what is left.

There is also a point about buying. The MIT finding that purchased tools succeed more often than internal builds is a case for buying well, not for buying instead of designing. A bought tool still lands in a workflow. Whether it succeeds depends on whether that workflow has an owner who can adapt the process around it, decide what the tool may do and capture what it saves. The purchase is the easy decision. The structure around it is the one that determines the return.

The right answer is not structure first and tools later. It is tools and structure in the same budget line. Deploy in a bounded domain, and in the same decision name the owner of the flow, change the measures, and install the controls. Let adoption lead the discovery. Do not let it lead the design.

Closing the gap: what leaders change, and in what order

Organisations do not operate according to strategy. They follow their incentives, shaped by capital and expressed through control.

That loop is worth stating precisely. Capital concentrates and defines where control resides. Control shapes how data is generated and used. Data drives execution. Execution produces outcomes, which attract further capital. Organisations scale what they are structurally designed to do, not what they intend to do. AI accelerates the loop. Weak structures fail faster. Aligned structures scale.

Closing the gap means intervening in that loop deliberately, in this order.

First, name the decisions, not the tools. Before scaling anything, list the decisions the AI will influence or make: approvals, prices, routes, responses, commitments. For each, name the person accountable for the outcome. If nobody can be named, the system is not ready to scale, however good the model.

Second, give one owner the whole flow. Assign end-to-end authority for the workflows where AI is meant to create value, including the budget and the right to change process, data and policy across functional lines. Without this, AI accelerates one step and the flow stays the same.

Third, re-cut incentives before you scale. Change what managers, teams and partners are rewarded for: outcomes, cycle time, quality and value captured, not headcount, budget size or effort billed. Renegotiate supplier contracts priced on effort. Do it before the savings appear, or the savings will be redeployed to protect the old measures.

Fourth, install control as infrastructure. Tier decisions by risk. Let low-risk automation move quickly. Keep human authority over high-consequence decisions. For every system, define what it may commit the organisation to, who can stop it and on what trigger, how affected people can seek review, and how its logic can be inspected. Put the CEO's name on the governance, not a committee's.

Fifth, fund the redesign in the same line as the licence. If the business case pays for the platform and not for the change to roles, measures and controls, the business case is for a tool. Make the operating model work part of the investment, with the same scrutiny.

Sixth, measure value captured, not adoption. Retire licences issued, active users and pilots launched as indicators of progress. Track the economic and operational outcomes the programme was meant to move, and report them to the board in the same format as any other capital investment.

Leaders also need to stop doing four things. Stop treating usage dashboards as evidence of transformation. Stop letting procurement and IT own AI strategy by default, because they own the purchase, not the operating model. Stop approving pilots that have no named production owner. And stop asking vendors what the operating model should be. They will answer honestly, and the answer will be their product.

Picture what this looks like in a mid-sized logistics company that wants AI to handle delivery exceptions. Done the tool-led way, it licenses an exceptions platform, connects it to the transport system and reports adoption. Done the structural way, the first month goes on listing the decisions involved: rebooking, refunds, carrier penalties, customer credits. Each gets an owner and a limit. One executive is given authority over the exceptions flow from first alert to customer resolution, across operations, finance and customer service. Team measures shift from tickets closed to exceptions resolved within the promised window. The system may rebook and issue credits below a threshold. Above it, a named person decides, and every automated action is logged for weekly review. Then the platform is deployed. The software is identical in both versions. Only the second one changes the business.

The sequence matters because each step creates the conditions for the next. Named decisions make ownership possible. Ownership makes incentive change enforceable. Incentives make controls credible. Controls make it safe to fund scale. The 7:1 perspective argues the same ratio from the deployment side: most of the work is governance and authority, and a small part is new execution capability. Why AI Transformation Fails traces what happens when that ratio is inverted.

Technology defines what is possible. Structure determines what becomes real.

Most organisations will adopt the technology. Few will adapt the structure.

What to do about it

- **Name the decisions, not the tools.** List the decisions AI will influence or make and name the person accountable for each outcome. If nobody can be named, the system is not ready to scale.
- **Give one owner the whole flow.** Assign end-to-end authority, including budget and the right to change process, data and policy across functions, for each workflow where AI is meant to create value.
- **Re-cut incentives before scaling.** Reward outcomes, cycle time, quality and value captured instead of headcount, budget size or effort billed, and renegotiate effort-priced supplier contracts.
- **Install control as infrastructure.** Tier decisions by risk and define for every system what it may commit the organisation to, who can stop it, how affected people seek review and how its logic is inspected.
- **Fund the redesign with the licence.** Put the change to roles, measures and controls in the same business case as the platform, with the same scrutiny.
- **Measure value captured, not adoption.** Retire usage dashboards as progress indicators and report AI outcomes to the board like any other capital investment.

Sources

- McKinsey & Company, The state of AI: How organizations are rewiring to capture value, March 2025: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-how-organizations-are-rewiring-to-capture-value>
- Fortune, MIT report: 95% of generative AI pilots at companies are failing, August 2025: <https://fortune.com/2025/08/18/mit-report-95-percent-generative-ai-pilots-at-companies-failing-cfo/>
- Microsoft and LinkedIn, 2024 Work Trend Index on the state of AI at work, May 2024: <https://news.microsoft.com/source/2024/05/08/microsoft-and-linkedin-release-the-2024-work-trend-index-on-the-state-of-ai-at-work/>

-
- BCG, AI Adoption in 2024: 74% of Companies Struggle to Achieve and Scale Value, October 2024: <https://www.bcg.com/press/24october2024-ai-adoption-in-2024-74-of-companies-struggle-to-achieve-and-scale-value>
 - McCarthy Tétrault TechLex, analysis of Moffatt v. Air Canada, 2024 BCCRT 149, 2024: <https://www.mccarthy.ca/en/insights/blogs/techlex/moffatt-v-air-canada-misrepresentation-ai-chatbot>
 - Zillow Group, Form 8-K on the wind-down of Zillow Offers, November 2021: <https://www.sec.gov/Archives/edgar/data/1617640/000161764021000085/z-20211102.htm>
 - Royal Commission into the Robodebt Scheme, Report and recommendations, July 2023: <https://robodebt.royalcommission.gov.au/publications/report>

Read online: bdgadvisory.com/perspectives/technology-innovation-operating-model-gap