ECONOMIST ENTERPRISE

Making AI deliver

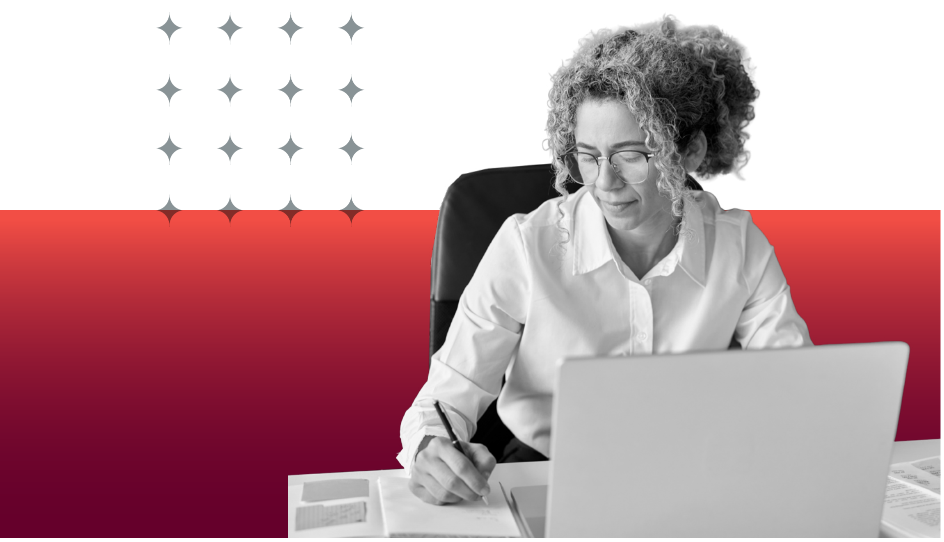
A benchmarking framework on how leading companies operationalise AI for impact

Supported by databricks

Contents

About this report

This Economist Enterprise report, supported by Databricks, examines how firms are turning advanced artificial intelligence (AI), including generative AI and AI agents, into durable operational capability. It is based on both qualitative and quantitative research components. One is an extensive survey of more than 1,200 respondents across regions and industries (see details below). Another component is a series of in-depth interviews with senior technology and data leaders (see acknowledgements below). An additional component is a roundtable with another set of senior technology and data leaders (see acknowledgements below). The final research component of the report is extensive desk research on existing work on the subject of operationalising AI in 2026. The analysis of the report is underpinned by a benchmarking framework designed to show how leading companies make AI work in practice.



Acknowledgements

We would like to thank the following executives for participating in interviews for this report and sharing their insights:

- **Ashish Agrawal**, chief information officer, **KONE**
- **Angie Ruan**, chief technology officer of capital-access platforms, **Nasdaq**
- **Annabelle Gerard**, vice-president for AI and data, **Stellantis**
- **Chas Murphy,** senior vice-president for direct-to-consumer data and analytics, **Disney**
- **Craig Price**, head of artificial intelligence and data science, **Suncorp Group**
- **Daniel Eitler**, chief AI and data officer, **Mercedes-Benz Group**
- **Gabriele Ricci**, chief data and technology officer, **Takeda**
- **Jose Manuel Silva**, vice-president for technology and chief digital officer, **Natura**
- **Karthik Iyer**, group vice-president and transformation leader for merchandising technology and AI, **Albertsons Companies**
- **Kaynaz Behdin**, senior vice-president for global digital transformation, data and AI, **Stellantis**
- **Maria Macuare**, senior vice-president and global chief data officer, **Mondelēz International**
- **Shimizu Seita**, managing executive officer, chief digital officer and chief executive of Solution Co-Creation division, **Tokyo Gas**
- **Swamy Seetharaman**, AI enabler, **CRED**
- **Tal Saraf**, chief information officer and senior vice-president for engineering, **Atlassian**
- **Vijay Mehta**, executive vice-president and general manager for global solutions and analytics, **Experian**
- **Wendy Batchelder**, senior vice-president and chief data officer, **Centene**
- **Youngjin Kim**, chief technology officer, **Yanolja Consumer Platform and NOL Universe**

We would like to thank the following executives for participating in a roundtable for this report and sharing their insights:

- **Brian Bischoff**, chief technology officer, **CapTech Ventures**
- **Sowmya Gottipati**, corporate chief information officer, **Estée Lauder**
- **Kim Hales**, chief data and AI officer, **NRG Energy**
- **Anand Mishra**, chief information officer, **MIO Partners**
- **Jason Sharples**, senior vice-president and chief information officer for global merchant and network services, **American Express**
- **David Ramirez**, chief information security officer, **Broadridge Financial Solutions**
- **Francis Shanahan**, chief technology officer, **Peloton**
- **Ben Celebicic**, chief product and technology officer, **Hinge**

Survey details

This study's global survey, developed and conducted by Economist Enterprise and supported by Databricks, polled 1,221 respondents. It was fielded between November 2025 and January 2026. Respondents work in large enterprises (with annual revenue of US$500m or greater), located in 18 countries in the Americas, Europe and Asia-Pacific.

The sample comprised:

Job titles: chief information officer (CIO) or equivalent, chief IT/technology officer (CTO) or equivalent, chief data/analytics/AI officer or equivalent, chief data scientist or equivalent, chief enterprise/data architect or equivalent, senior vice-president/vice-president of IT/data/engineering/AI or equivalent

Locations: Australia, Brazil, Canada, Denmark, Finland, France, Germany, India, Italy, Japan, Mexico, Netherlands, Norway, South Korea, Spain, Sweden, United Kingdom, United States

Industries: financial services, banking and insurance; healthcare, pharmaceuticals and life sciences; retail and consumer goods; manufacturing and automotive; media and entertainment; energy, oil and gas; telecommunications; digital natives

Executive summary

The gap between what artificial intelligence (AI) can do in 2026 and what firms actually use it for is vast. To understand why, we surveyed 1,200 executives at firms using advanced AI and spoke to 25 senior technology leaders across different industries. Our conclusions are sobering for optimists and encouraging for pragmatists. We find a corporate world awash with AI activity yet short of the impact that boosters promise. At the same time, we describe how firms that have clear plans for how to use AI and have invested in digital plumbing and their staff are pulling ahead fast.

This report examines the mechanics of success with AI. It introduces a benchmarking framework spanning nine distinct capacities, from strategy and data foundations to governance, work redesign and the demands of autonomous agents. Our framework recognises that AI capability does not accumulate evenly. One company can have superb AI infrastructure but blunt the technology with feeble change management. Another can deploy agents at speed while lacking the governance to keep them honest. Six things stand out.

1. Activity is not impact
In 2026, high levels of AI deployment mask thin returns. Our survey finds that more than four in five executives say their AI programmes are beating expectations. Yet only about two in five firms formally require teams to track business impact, from cost savings to revenue and efficiency. Proximity to the work narrows the distance between enthusiasm and evidence. Nearly nine in ten chief technology officers say their AI roll-out is ahead of schedule, but only three in four senior vice-presidents agree. Firms that confuse a growing list of AI experiments for progress will find their pipelines for scaling the technology clogged and their boards impatient.

2. Data plumbing is the binding cost
We find that just over half of firms with unified data architectures cite data storage, movement and duplication as their biggest ongoing AI expense, rising to roughly two-thirds among firms with less integrated environments. In either case, that share is more than double the proportion pointing to computing infrastructure. The gap between building a prototype on clean data and running AI at enterprise scale is, as one executive puts it, "not even in the same universe". Firms that consolidated their data estates before the generative-AI wave arrived are pulling ahead. Those still patching fragmented systems find that every new AI use case lays bare weaknesses they could previously ignore.

3. Firms struggle to move from pilot to production

About three in five firms take between 7 and 12 months to ship an AI project. A similar share of firms either lack a fully established development life cycle for AI projects or do not apply it consistently. The result is perpetual piloting in which AI experiments multiply without the discipline to scale what works or kill what does not. The firms breaking this pattern share three features: a structured life cycle, disciplined attrition and design for reuse. Together, they form what our framework calls the scaling engine.

4. Governance thins out where it matters most

About three in five firms review AI systems during development and before deployment. Fewer than two in five continue that oversight after a system goes live—the stage where AI models drift, data shifts and edge cases multiply. Worse, one in eight firms reviews governance only when something goes wrong. Companies that govern AI across its full life cycle, from inception through live operation, suffer fewer failures and scale with greater confidence.

5. Culture, not technology, most often determines whether AI scales

The executives we interviewed returned, almost without exception, to the same point. The hardest part of making AI work is not building the models but rewiring the organisation around them. Task-level job redesign, meaningful training and the right incentives matter more than the sophistication of AI systems. Our survey captures the imbalance. Half of firms cite human review as a top ongoing cost, yet only 4% point to employee upskilling. Firms are wrong to think that they can keep AI running without investing in the people who must work alongside it.

6. Using agents requires fixing existing weaknesses

The headline numbers about AI agents are striking. About three in five leading AI adopters now have autonomous systems doing real work. But the governance structures that should accompany them lag well behind. Fewer than half of firms mandate a formal framework for autonomous systems. The barriers to scaling agents, from accuracy to integration with legacy systems, are familiar from earlier AI, but they bite harder when software acts rather than merely advises. The firms furthest ahead invested in the control layer of AI before they invested in autonomy. That sequence matters for any firm vying for success.

Introduction

Now make it deliver

In 1987 Robert Solow, a Nobel-prizewinning economist, quipped that one could "see the computer age everywhere but in the productivity statistics". Four decades later, many business leaders feel a profound déjà vu. AI boosters say the technology is revolutionising the global economy, yet the productivity miracle they promise remains stubbornly elusive. America's Bureau of Labor Statistics finds that in 2025 labour productivity in the business sector rose by 2.2%, only slightly above the long-run average of about 2%.[1] As 2026 unfolds, the transformation of the corporate world has yet to happen.

This is not for want of technological firepower. AI models that in 2022 struggled to summarise longer PDFs have been succeeded by models that can reason across files, write and debug code, interpret images and—in their incarnation as agents—complete multi-step tasks with minimal human oversight.

Nor is it for a lack of enthusiasm. Generative AI has crept into most corporate departments because its ability to understand ordinary language makes it easy to use. Earlier types of AI, such as machine-learning systems that sift vast data sets, remain powerful but usually require expensive experts to tailor them to a specific task, which limits their adoption.

What bosses now face is a daunting set of practical challenges. They must confront the economics of running models at scale and figure out how to measure profits from AI. Governance looms, too. Firms need to establish how to police harmful AI use and who is accountable when AI agents make costly mistakes. Most troubling of all is a widening gap between the might of AI systems and how ill-prepared some organisations and their staff are to use them meaningfully.

This report examines how companies close that gap. Drawing on a survey of 1,200 executives at organisations already using advanced AI, conversations with 25 senior technology leaders and desk research, it examines what meaningful AI use looks like in 2026. It includes a benchmarking framework designed to separate firms that have woven AI into the fabric of their business from those still running impressive experiments at the edges.

1 U.S. Bureau of Labor Statistics. Fourth Quarter and Annual Averages 2025, Revised. March 2026. Available at: https://www.bls.gov/news.release/pdf/prod2.pdf

Beware of high activity

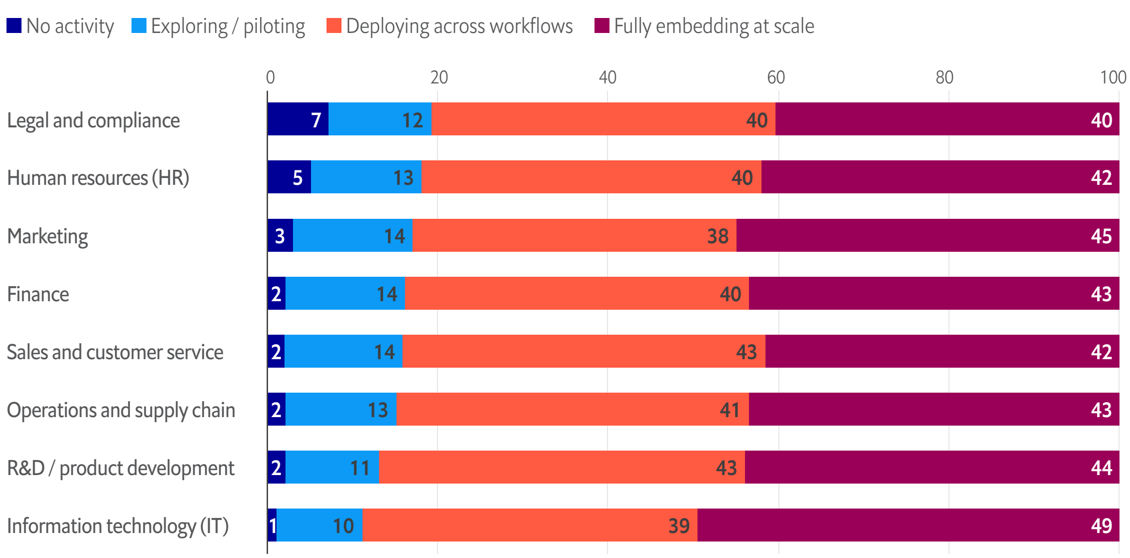
Step into most firms today and AI seems to be everywhere. A year or two ago, companies urged staff to experiment with the technology and launch as many pilots as they could. It worked. In our survey of leading adopters, between eight and nine in ten firms are deploying AI, or have fully embedded it, across every business function we asked about (see Figure 1). IT leads; marketing, research and development, and sales are nearly as advanced. Even legal and compliance, the laggard, sees AI used at scale by about two in five firms.

But high activity does not mean high impact. Our survey exposes a gap between corporate confidence and measured results. More than four in five executives boast that their overall returns from AI beat expectations so far. Yet fewer than half say that their firms actually require teams to track these gains.

"Our ambition is not AI adoption," says Jose Manuel Silva, vice-president for technology and chief digital officer at Natura, a Brazilian cosmetics group. Adoption is an activity metric, he argues. "What matters is 'kinetic value'—the moment an AI model converts latent capability into measurable business outcomes." Gabriele Ricci, chief data and technology officer at Takeda, a global pharmaceutical company headquartered in Japan, puts it more bluntly: "Our board is not interested in the number of AI pilots or prompts any more. They are focused on the impact of AI on the profit-and-loss account."

Figure 1: The mirage of AI everywhere

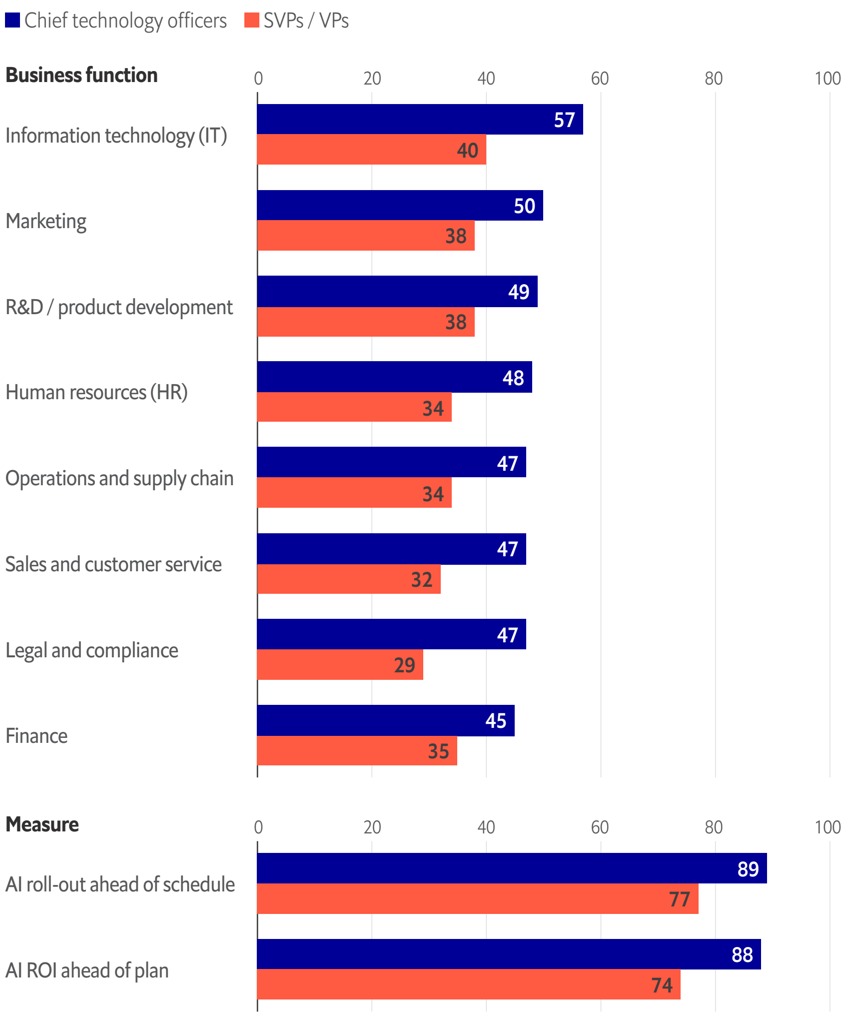
Stage of advanced-AI adoption by business function, % of respondents



Source: Economist Enterprise survey

Figure 2: Mind the gap

Perception of AI integration, chief technology officers vs (senior) vice-presidents



Source: Economist Enterprise survey

The gap between confidence and measurement shows up most starkly in a split by seniority. Nearly nine in ten chief technology officers say their AI roll-out is ahead of schedule and delivering strong returns on investment. Vice-presidents—the managers responsible for making AI work in practice—are markedly less sure: only about three in four agree. The most concrete measure of that divide is whether AI is actually embedded at scale. In the IT space, about three in five technology chiefs say it is, against only about two in five vice-presidents (see Figure 2). In legal and compliance, the gap is wider still—nearly half of chief technology officers cite full integration, against fewer than three in ten vice-presidents. Those setting strategy see tools deployed. Those running operations see how little has changed.

The main reason why activity has outrun impact is that deploying a tool is far faster than changing how an organisation works around it. To turn AI from an impressive novelty into a dependable system, firms must redesign workflows, retrain staff and rebuild the processes through which decisions actually get made. A survey from 2025 of more than 1,250 executives by Boston Consulting Group, a consultancy, finds that just 5% of firms realise AI's value at scale, achieving about five times the revenue growth of their peers.[2] About three in five, despite heavy spending, report no material return at all.

2 Boston Consulting Group. Are You Generating Value from AI? The Widening Gap. January 2025. Available at: https://www.bcg.com/publications/2025/are-you-generating-value-from-ai-the-widening-gap

The end of holiday economics

For a while, AI felt almost free. The first wave of tools arrived with enough capability to impress and enough ease of use to spread without formal budget lines. Pilots were cheap, experiments multiplied and the question of what AI cost, in both money and effort, rarely came up.

The shift to scaled deployment has exposed the gap between activity and return. About four in five firms say their AI initiatives are backed by strong business cases with expected returns on investment. Yet only about two in five formally require teams to measure business impact, whether cost savings, revenue gains or efficiency improvements.

The problem starts with what is being counted. Ashish Agrawal, chief information officer at KONE, a global elevator and escalator company, describes the dominant metric as "return on employee", the notion that AI tools make workers more productive. He argues that this is "very hard to measure" and harder still to connect to higher operational profit. Time saved is only value created if it flows into something the business can count, from lower operating costs and faster decisions to higher revenue.

The measurement problem has a cultural dimension, too. Framing AI's value as a reduction in headcount is not only analytically imprecise (a 10% to 20% efficiency gain in a given workflow rarely justifies a staffing cut) but also tends to generate resistance that derails adoption. Organisations that do not provide assurances against job losses are likely to see lower uptake of AI tools than those that foreground freed capacity.

The cost side compounds the difficulty. What most firms spend money on when they run AI at scale turns out to be different from the computing and licensing fees that dominate public debate. The real burden, we find, falls on moving and maintaining data, and on the staff hours required to check what algorithms produce. The models themselves, rapidly commoditising and falling in price, are the least of it (see Figure 2.1 in chapter two).

This reframes investment decisions. Swamy Seetharaman, who leads AI enablement at CRED, an Indian fintech, draws a useful distinction: firms should treat frontier models as commodity inputs but invest in the internal systems that translate intelligence into decisions, assets that compound over time. The pressure to make this case rigorously is intensifying. Kaynaz Behdin, senior vice-president for global digital transformation, data and AI at Stellantis, a carmaker (whose biggest shareholder, Exor, part-owns Economist Enterprise's parent company), sets a clear threshold: AI projects will be funded if they deliver measurable value in 12 to 18 months (on metrics such as efficiency gains, cycle-time reduction or operational impact). That is demanding given that most firms in our survey take between 7 and 12 months simply to move a project into live production, meaning returns must materialise from the moment something goes live.

Not everyone chases that clock. Shimizu Seita, chief digital officer at Tokyo Gas, manages AI expectations at the executive level, creating awareness that the returns from the technology may take time to materialise as the utility rewires its processes around the technology. That patience may be better economics than it appears (see chapter five). Our survey offers a hint: among very large firms, only about three in four say their AI initiatives are backed by strong business cases, compared with about four in five among their smaller peers. Complexity scales faster than confidence.

The Ferrari without a driver

There is a cruel irony to labour-saving automation: human beings invariably stand in its way. Up until 2023, the bottleneck for AI was the sheer cost of computing power and the crudeness of the algorithms. Today the models are ready but the workforce rarely is. Mr Silva, head of technology at Natura, reaches for a vivid image: "I have a Ferrari, but I don't have the driver."

Mr Silva explains that teams in Natura's Pay technology division adapted quickly because they "were born in data", so staff were already comfortable with hypothesis testing and feedback loops. But in departments like human resources and logistics, the bottleneck was never the software. It was skill, mindset and the cultural permission to experiment, says Mr Silva. As a result, businesses often move at two speeds: algorithms spread rapidly where the culture welcomes them and stall completely where it does not.

Leading firms respond by investing in AI training. Suncorp, an Australian insurer, has partnered with the University of Sydney to deliver an "AI fluency" course for senior leaders. Craig Price, the company's head of AI and data science, says that "hopes and dreams have to be converted into real actions and it's critical that our leadership understands how AI works and can benefit their part of the organisation."

Our survey captures the shortfall. About three-quarters of respondents say they have redesigned job descriptions to reflect AI. Yet only 4% name employee upskilling as a significant ongoing cost. Evidently many firms are rewriting titles without rebuilding roles. Chapters three and five explore in detail why closing that gap is one of the hardest tasks in 2026.

Benchmarking framework

A common way to gauge AI progress inside a company is to count everything from software licences to test pilots and tools in use, then slot the results into a few steps of upward progression towards "AI maturity". This is why most existing frameworks assessing how firms are doing with AI arrange them on a neat staircase: experimenting, piloting, scaling, optimising. That makes for a clean slide deck but a poor map of reality.

In practice, AI capability does not accumulate in a straight line. A company can have superb data infrastructure and feeble change management. It can deploy agents at speed while lacking the governance to keep them honest. It can redesign job descriptions on paper without altering a single workflow in practice. Linear frameworks hide these mismatches.

The benchmarking framework at the heart of this report measures what organisations can reliably do across nine distinct capacities without forcing them onto a single prescribed path. Those capacities span the full terrain of making AI work: from strategy and value discipline, through technical foundations, scaling engines, governance and control, and the question of whether AI is truly built into real work or bolted on beside it. They also include what many frameworks relegate to an afterthought: work redesign and skills, democratisation with guardrails and the operating model underpinning AI. Finally, the framework considers the specific demands of agentic AI.

Capacity	What it covers	Yardsticks of strength
Strategy and value discipline	Whether AI efforts serve the business strategy, are run with financial accountability and produce clear outcomes. This capacity acts as a filter: without it, the pipeline fills up with projects that cannot be evaluated or stopped.	- AI goals are expressed in the language of business outcomes, not technology capabilities - Work is funded, stopped or scaled based on measured results - Business cases include both building and running costs - Outcomes link to company-wide goals and shape the next wave of investment
Technical foundations	Data architecture, platform standards, integration infrastructure and the plumbing that connects AI models to the work they are meant to perform. This is the enabling layer on which almost everything else depends.	- Data is consolidated into a form that is findable, accessible, interoperable and reusable (FAIR) - Common standards apply across units, with controlled exceptions - Lineage and metadata are maintained for audits and root-cause work - Reusable, secure integration reduces bespoke builds and duplication

Capacity	What it covers	Yardsticks of strength
Scaling engine	How consistently promising experiments become monitored, governed services in daily use. This capacity separates firms that can repeat success from those that run impressive one-off pilots.	- A structured life cycle governs the path of AI projects from idea to production - Evaluation and monitoring are routine once systems are live - Projects are killed when they fail to meet the bar - Incidents are handled with defined procedures and drive improvements
Built into real work	Whether AI sits inside core workflows and products, rather than beside them as a bolt-on tool.	- AI is embedded in the tools people already use - Hand-offs, exceptions and approvals are formalised - Adoption is sustained and measured - Workers and end-users help design the system
Governance and control	Risk, compliance and oversight across the full AI life cycle, with separate consideration for autonomous systems. Governance that covers only the approval stage and thins out after deployment creates a false sense of security.	- A clear governance framework spans design, development, deployment and live operation - Actions and outputs are traceable - Policies evolve with evidence from outcomes and incidents - Escalation and kill-switch norms are clear and used
Work redesign and skills	Whether roles, tasks and training are being rebuilt around AI. This includes the hard work of separating what machines can do from what people must judge, and investing in the skills to make the collaboration productive.	- Clear task boundaries define what is human-led, AI-led and collaborative - Meaningful, role-specific training is funded and delivered - Redesigned roles show measurable gains in output or quality - Incentives reward effective use of AI, not merely adoption
Democratisation	How broadly non-technical staff can use AI and data safely, with the right support. The goal is broad access without ungoverned sprawl.	- Many workers can use AI without creating duplication or security risk - Guardrails are simple, enforced and understood - Training, support and community are in place - Self-service tools put data and AI in the hands of those closest to the work
Operating model and ecosystem	How AI work is organised, funded and governed across teams and suppliers. This is the organisational infrastructure behind every other capacity, though it is often left implicit.	- Who decides, builds and runs AI is explicit and understood - Product and platform teams are stable, not assembled ad hoc - Incentives reward reliable scaling, not heroic one-off launches - Relationships with outside vendors are managed centrally
Agentic AI	Whether a firm can deploy autonomous AI systems that act with limited human oversight and govern them once they are live. This capacity depends on every other and amplifies any weakness elsewhere.	- Autonomy levels and permitted actions are defined for each agent - Approvals and exception paths exist and are used - Agent actions are observable and auditable - Simulation and testing are standard before higher-stakes deployment

Some capacities in our framework enable others. Strong technical foundations, for example, are a prerequisite for deploying autonomous agents. Robust governance enables speed rather than blocking it, because teams that trust their controls experiment more freely. But weaknesses in one area can cap progress elsewhere. A firm with brilliant platforms but no process for killing failing projects will clog its pipeline. A firm that automates aggressively before its governance is ready creates fragility. And a firm that redesigns jobs without investing in training produces anxiety rather than adoption. These dependencies and caps explain why organisations that look advanced in one respect can remain stuck in another.

Our framework helps draw out distinct organisational types. Some firms are platform-first scalers: strong on infrastructure and tooling but struggling to embed AI in daily work. Others are governance-led integrators: disciplined and controlled but slower to experiment and scale. A third group are autonomy pioneers: racing ahead with agents and automation, but outrunning their governance. Each path has strengths and blind spots and none is inherently superior. What matters is whether a firm understands the shape of its own capability—and what it should do next.

Two design choices deserve brief explanation. First, our framework treats agentic AI as a separate capacity rather than merely a more advanced form of the same technology. Autonomous systems that act without a human checking every step create qualitatively different demands for governance, auditability and trust. Second, the framework places workforce redesign at its centre. This reflects our conversations with executives across sectors who kept coming back to the idea that in 2026 it is culture, not technology, that most often determines whether AI scales or stalls.

Throughout this report, our framework organises our judgments and sharpens our comparisons. We surface it in cut-out boxes that link our chapters with its metrics. The goal is to give leaders a practical way to see where they are strong, where they are weak and what they should do next.

Chapter 1: Strategy headstart

Strategy, priorities and the value case

Corporate history is littered with the corpses of firms that mistook a new technology for a strategy. During the dotcom boom of the late 1990s, many businesses assumed a web address guaranteed riches, only for outfits like Webvan, a pioneering online grocer, to collapse spectacularly. Established retailers discovered the hard way that online sales incurred new variable costs for fulfilment while the fixed costs of their physical shops remained.

AI presents a similar temptation. Renting computing power and launching scattered pilots is easy, especially as AI models become rapidly commoditised. But doing so can also lead to a paralysing glut of not-so-meaningful AI projects that sink corporate energy but deliver little value.

The companies making genuine progress have answered the question of strategy before they started. They connect the technology to specific business outcomes, build frameworks to measure whether those outcomes are achieved, and give leadership the tools to choose where to invest and where to stop. Strategic clarity is the earliest condition for making AI useful—and among the hardest to maintain as the technology rapidly evolves and spreads.

Benchmark capacity in this chapter:

Strategy and value discipline

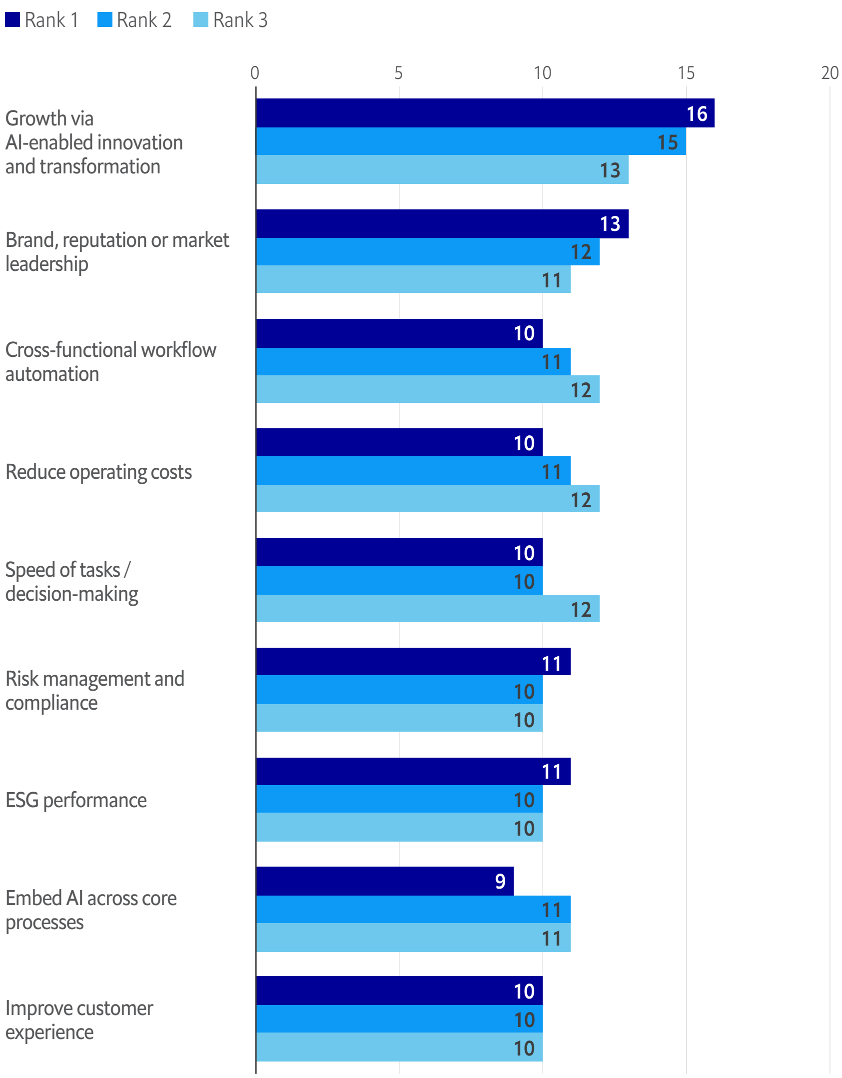
Strategy and value discipline is the capacity that tests whether AI efforts serve the business—the most foundational of the nine. It measures whether a firm's AI goals actually shape decisions: which AI projects are pursued, which ones are stopped, and whether the connection between AI activity and business outcomes is enforced.

Firms strong in this capacity can point to a live link between AI spending and the profit-and-loss account. Firms weak in it often have board-level ambitions and an unmanaged pipeline below them.

The cost of getting this capacity wrong is high. Without value discipline, a firm's AI portfolio fills with pilots that cannot be evaluated and projects that cannot be stopped. That clogs the scaling engine and blunts every other capacity in the framework.

Figure 1.1: A matter of growth

Top goals for advanced-AI investment over the next two years, % of respondents*



Source: Economist Enterprise survey
* Respondents were asked to rank their top three goals

A means to an end

Every sweeping technology, from cloud to 5G, spawns its own corporate vocabulary, and AI is no exception. Managers talk about "AI-first" strategies, "AI transformation" and "AI-native" workflows as though the technology were itself the objective. The firms furthest ahead have resisted this linguistic drift. They define AI's role not by what the technology can do but by what the business needs done.

Angie Ruan, chief technology officer of capital-access platforms at Nasdaq, a financial-technology platform for global capital markets, frames it plainly: "AI is an accelerator to the flywheel we already have." Nasdaq embeds this subordination in its internal language, distinguishing between "AI on the business"—improving internal operations—and "AI in the product"—serving clients better.

Mercedes-Benz approaches the technology similarly. Daniel Eitler, its chief AI and data officer, says the German carmaker aims to build the world's most desirable cars, cost-efficiently and with the best people—all "accelerated by AI and data". The firm anchors AI to goals that existed before any large language model was trained, giving Mr Eitler a mandate that is easily understood across the entire organisation.

Our survey confirms that this growth-oriented framing dominates across leading AI adopters. When asked to rank their top goals for AI investment over the next two years, executives place growth through AI-enabled innovation and transformation first (see Figure 1.1). Cost reduction, long the default justification for automation, comes only fourth. The pattern holds across company size, though with telling variation by sector: manufacturers and automotive firms, for example, place speed of decision-making higher than other industries do, while financial-services companies are more likely to emphasise risk management and compliance.

But ambition is not operational reality. Our survey tests whether firms actually wire their AI initiatives to company-wide objectives. Just under half of all respondents say that linking AI outcomes to broader goals—such as customer experience, growth or risk reduction—is formally required for all relevant teams. A further two in five say they usually do it, but without a mandate. That leaves about one in seven firms where the connection between an AI project and a corporate goal is made only occasionally, or not at all. The gap between declaring AI a strategic priority and enforcing that priority by linking the two remains wider than the headline confidence would suggest.

Sorting the haystack

The spread of priorities matters as much as their ranking. No single goal dominates overwhelmingly, suggesting that most firms pursue AI along several fronts at once. Maria Macuare, senior vice-president and chief data officer at Mondelēz International, a snack-food giant whose brands include Oreo and Cadbury, divides AI's role into three "buckets": employee productivity, AI-infused existing software platforms and a handful of "game changers"—concrete use cases where AI can materially shift business outcomes.

The question facing all companies is who identifies these priorities—and how. Some take a two-track approach that blends direction from the top with ideas generated by the people closest to daily work. At Mercedes-Benz, a network of "AI officers" inside each business division hunts for everyday tasks and processes that AI could improve and feeds these ideas back to the centre, which, in turn, ensures that different departments are not inventing the same tools twice.

Takeda takes a more centralised approach, reflecting both pharma's heavy regulation and the firm's consensus-driven culture. It governs digital investment through a "digital portfolio committee" that Gabriele Ricci, its chief data and technology officer, co-chairs with the chief executive. The committee ranks AI projects with the same discipline as Takeda's drug pipeline: a two-stage path from idea to minimum viable product, then to enterprise scale. "We are probably the only pharma company that prioritises our scientific pipeline and our digital pipeline with the same process," Mr Ricci says.

Suncorp, by contrast, begins at the edges and filters inwards. Craig Price, who leads the firm's AI and data-science practice, started with 120 ideas gathered across the business, then themed, consolidated and ranked them. Only then did a central steering committee—which "included representatives from across the business", he says—decide which initiatives to back and what returns cleared the bar.

Capacity box 1

Strategy and value discipline

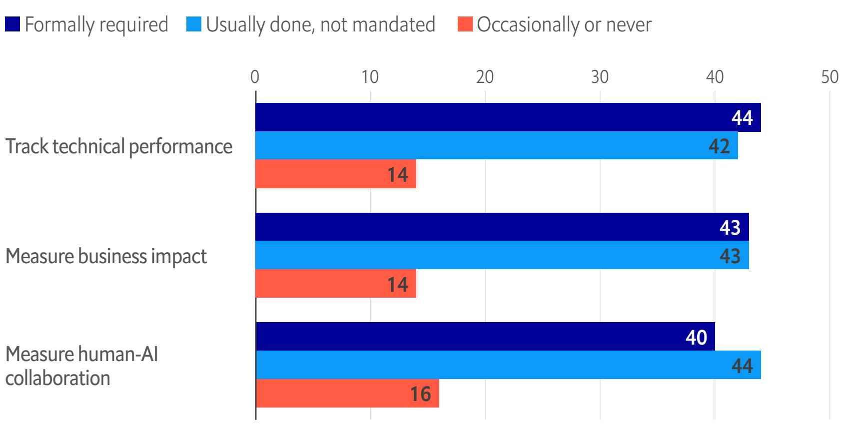
Firms approach the connection between AI and business strategy in broadly three ways. The first group treats AI as a corporate-level accelerator. These firms embed AI explicitly within their overall business strategy as a means of reaching goals that predated the technology. Nasdaq's distinction between "AI on the business" and "AI in the product", and Mercedes-Benz's formulation that the company will build desirable cars "accelerated by AI and data", both exemplify this pattern. The yardstick of strength is whether the language of AI and the language of business goals are genuinely fused.

The second group prioritises AI without anchoring it. These firms invest heavily and report high confidence, but lack the mechanisms to connect AI work to corporate outcomes. Our survey captures this gap: four in five firms say their initiatives are backed by strong business cases, yet fewer than half formally require teams to track impact. The risk is a swelling portfolio of AI experiments that cannot justify their cost.

The third group pursues AI reactively, responding to competitive pressure or executive enthusiasm without a clear hierarchy of priorities. These firms tend to generate the largest backlogs of use cases and the highest rates of project abandonment. Several of the executives we spoke to described this pattern in their own organisations before a deliberate reset.

Figure 1.2: Forgot how to count?

AI-evaluation practices across organisations, % of respondents



Source: Economist Enterprise survey

Despite their structural differences, neither approach leaves AI strategy to chance. Firms succeed when they treat AI projects as deliberate bets, guided jointly by technologists and business managers. The effect is to compress the distance between ambition and accountability—the dependency that most reliably separates firms with real value discipline from those whose pipelines grow without limits (see Capacity box 1).

Measuring what matters

When sky-high corporate confidence in AI has yet to yield hard financial returns, firms need to figure out how to track exactly what the technology achieves. Our survey finds that many are trying to do so, though their rigour varies. About two in five executives say that measuring business impact, from cost savings to revenue gains and efficiency, is formally required (see Figure 1.2). A further two in five say it is usually done but not mandated. The numbers are similar for tracking technical performance such as accuracy, reliability and error rates.

"I get a lot of people telling me that many millions of lines of code are being generated by AI, and I ask them: what did this materially change for KONE?"

Ashish Agrawal, chief information officer, KONE

The picture grows patchier for the human side. Measuring how effectively people and AI work together, from changes in productivity to decision quality and employee trust, is required by two in five and done informally by a similar share, but one in six admit it happens only occasionally or not at all.

The most advanced firms define success in terms all employees can grasp. Wendy Batchelder, chief data officer at Centene, an American health insurer, says that the firm uses the technology to speed up access to care and improve the health of the communities it serves. "AI does not determine whether care is denied," she says. "It's giving you a faster path to yes."

Suncorp looks closely at the human impact. Mr Price says the Australian insurer counts the time its technology saves. A single tool for call-centre workers, for instance, has spared them more than 30,000 hours of labour. But bosses also measure whether customers are happier, and whether staff trust the software enough to resolve problems without asking a human expert. To do this, Suncorp surveys its workers to establish how they feel about the new tools.

Ashish Agrawal, chief information officer at KONE, distils the point. When his team tells him that AI has written millions of lines of code, he asks: "What did this materially change for KONE?" He refuses to celebrate until his team can prove that the same number of developers now produce more, or that fewer are needed to do the exact same work. "I need those hard metrics defined," he says.

The accountability test

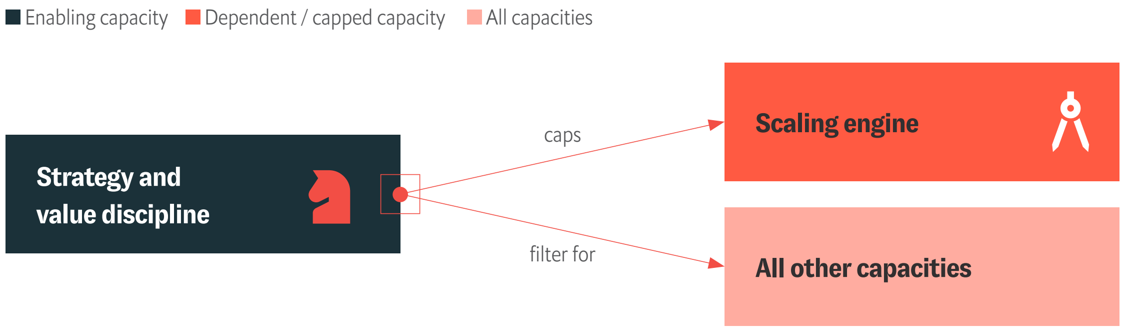
The firms with the strongest strategy and value discipline share three traits. They define AI's role in the language of business goals, not technology capabilities. They build mechanisms, from portfolio reviews to financial thresholds and board-level oversight, that require every AI project to justify its existence in terms the profit-and-loss account can verify. And they create a mechanism to stop work that fails to meet the bar.

What separates leading firms is not ambition—nearly every firm we surveyed is ambitious about AI. What matters is accountability. Firms that require teams to track impact, link outcomes to goals and apply lessons learned across their AI projects turn ambition into discipline. Laggards tend to confuse a growing list of experiments for progress.

Weak value discipline clogs the pipeline, obscures real gains and makes every other capacity harder to evaluate. It is the filter through which all other capabilities must pass.

Dependencies and caps

Chapter 1: Strategy and value discipline



Industry profile 1

The Hippocratic algorithm

One of the first things medical students are taught is to do no harm. Computer-science students are often told that their industry is about shaking things up. As AI enters healthcare, these two philosophies are colliding. Although pharmaceutical companies such as Takeda are eagerly using the technology to discover new drugs, the broader healthcare sector is bogged down by fragmented data, strict privacy rules and the fear of legal liability. Yet the industry's appetite is evident in how executives want to use AI agents, the frontier of advanced AI in 2026. About half point to personalised agents that monitor patients remotely and adjust their care plans. AI-driven diagnostics and the management of clinical trials follow closely, each cited by about two in five healthcare executives.

Treating patients with AI, however, is far more complex than predicting consumer behaviour. Algorithms can inherit dangerous biases, and how they reached a diagnosis often cannot be explained. When asked to name their greatest challenge to deploying autonomous agents, healthcare bosses split almost evenly across three barriers. About a quarter cite patient safety and the rigorous burden of clinical validation. A similar share flag the difficulty of protecting sensitive information under state regulations. And about a quarter point to gaining the trust of clinicians.

Health systems increasingly govern AI based on the risks involved. A study from 2026 by researchers at the University of Chicago reviewed 35 frameworks for AI implementation in healthcare.[3] It found that the best models divide AI into distinct tiers. Diagnostic tools face strict safety checks, whereas administrative software receives a lighter touch. Health systems that treat a note-taking app with the same scrutiny as a diagnostic algorithm risk bogging down regulators and missing real dangers. This matters because the potential benefits of clinical AI are vast. A rare disease might appear only once in a doctor's career. Yet an AI trained on huge data sets can spot signs that no single clinician would reliably detect.

Doctors are most eager to use AI to reduce their paperwork. Many spend two to three hours each evening completing clinical notes, a regulatory requirement that they cannot delegate. But AI assistants can now listen to consultations in real time and draft records for the doctor to review. A study published in 2025 by researchers from Yale and several large American health systems surveyed more than 250 clinicians across six organisations, examining the impact of an AI tool for drafting medical notes.[4] It found that after a month of using an AI note-taker, the share of those reporting burnout fell from about half to roughly four in ten. Doctors spent less time on after-hours paperwork and improved how they interacted with patients.

3 Nature (npj Digital Medicine). Advancing healthcare AI governance through a comprehensive maturity model based on systematic review. February 2026. Available at: https://www.nature.com/articles/s41746-026-02418-7

4 Olson et al. Use of Ambient AI Scribes to Reduce Administrative Burden and Professional Burnout. October 2025. Available at: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2839542

Chapter 2: Below the waterline

Technical foundations: data, platforms and models

Much of the magical thinking that surrounds AI stems from the fact that its popular incarnations—chatbots—can be used by virtually anybody who can type. But for most large companies, deploying the technology remains an exercise in heavy digital engineering. That engineering is more consequential than the technology's problem-solver reputation suggests.

The divide between the firms pulling ahead and those stuck in endless experiments comes down to three things. First, the strength of their underlying data. Second, how firms structure their data and AI platforms—centrally, locally or as a blend of the two. And third, how firms choose and orchestrate their outside models as algorithms become cheap commodities. Getting all three right tests whether a company is willing to spend heavily on infrastructure invisible to customers.

Benchmark capacity in this chapter:

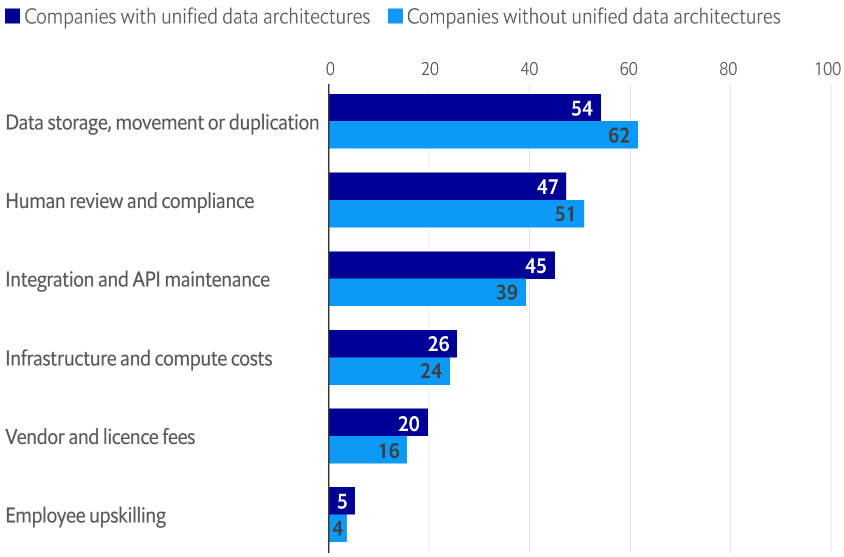
Technical foundations

The technical foundations of a firm are the enabling layer on which almost everything else depends. This capacity includes data architecture, platform standards, integration infrastructure and the plumbing that connects AI models to the work they are meant to perform.

The absence of strong technical foundations caps almost every other capacity in our benchmarking framework. Firms with weak foundations can run impressive pilots but they cannot scale them. They can get staff to use AI wisely but won't be able to give them AI systems that actually help. And they cannot safely pursue agentic deployment because it depends on data that is consistent, governable and lineage-tracked.

Figure 2.1: The hidden data tax

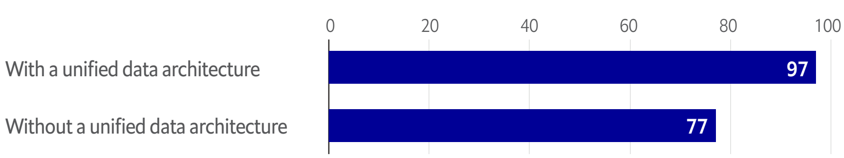
Most significant ongoing costs of running AI systems, % of respondents



Source: Economist Enterprise survey

Figure 2.2: The consolidation dividend

Share of firms reporting that AI spending is paying back faster than planned, by data-architecture-unification status, % of respondents



Source: Economist Enterprise survey

> **"If you can infuse AI on your data and it works, it means your data is really ready and follows the FAIR framework—findable, accessible, interoperable and reusable."**

Maria Macuare, senior vice-president and global chief data officer, Mondelēz International

Build now, or pay later

For more than two years, companies have raced to overhaul their systems for AI. Our survey finds that most are pleased with their work. About four-fifths of executives say that their data foundations are now strong. But this confidence may be hasty due to the demands of scaling AI. Many mistake the ease of building a prototype for the ability to run the technology at scale. Getting the data right for a single use case is "easier, faster and cheaper" than doing so "for all possible things", notes Maria Macuare, who leads data work at Mondelēz, a snackmaker. The two approaches, she adds, are "not even in the same universe".

Moving from a handful of isolated experiments to everyday execution exposes weaknesses in the digital pipelines that route information across a business. Each new AI use case forces firms to move more data between systems, creating endless copies that require documentation and tracking.

The bill is already arriving. Just over half of firms with unified data architectures cite data storage, movement and duplication as their biggest ongoing AI expense, rising to roughly two-thirds among firms with less integrated data architectures (see Figure 2.1). In either case, that share is more than double the proportion pointing to computing infrastructure.

The financial case for good data architecture is strong. We find that almost all firms that have unified their data architecture say that their AI spending is paying back faster than planned, compared with just over three-quarters of those that have not (see Figure 2.2). That gap makes consolidating data estates one of the most reliable predictors of whether AI investment will pay off.

Bad plumbing also imposes a strategic penalty. Flimsy infrastructure makes it impossible to run self-service platforms for employees, let alone deploy swarms of autonomous AI agents, without costs spiralling out of control. For Ms Macuare, AI is the acid test of a firm's plumbing.

Capacity box 2

Technical foundations

Of a firm's technical foundations, its data infrastructure matters most. On this count, firms divide into three broad camps. The first group consolidated early. These firms invested in unified data platforms, common standards and clean lineage before the generative-AI wave arrived. Craig Price, head of AI and data science at Suncorp, traces the insurer's readiness back more than a decade. Ashish Agrawal, chief information officer at KONE, describes how "achieving a common data platform on Databricks" allowed the firm to create reusable data products that "increase speed to market for APIs, digital services and machine learning" across its global operations.

The second group is consolidating under pressure. These firms began their data transformation in response to AI, and the work is not always complete. They can run single pilots on clean subsets of data, but scaling exposes fragmented systems, inconsistent definitions and missing lineage. Our survey's finding that nearly three in five firms cite data storage and movement as a top ongoing cost reflects this group's reality.

The third group has not yet started. These firms rely on department-level data stores with little central oversight. For them, even a successful pilot is difficult to replicate, because the data that powered it does not exist in a form other teams can use.

The dangers of ignoring data quality go beyond inefficiency. Wendy Batchelder, chief data officer at Centene, captures the risk. "If you have a faster train to the wrong station, that comes with additional risk," she says. In healthcare, where a bad inference can affect a patient's treatment, bad data amplifies errors.

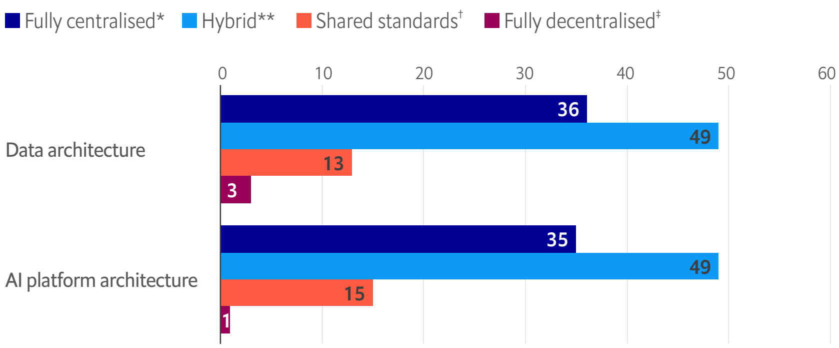
Consolidating data is painful but necessary. Consider the case of Natura. After three acquisitions and a merger, it was left with nine separate data lakes (centralised stores of raw information). The team led by Jose Manuel Silva, the firm's head of technology, merged those nine lakes into a single unified platform and consolidated more than 1,200 applications. He describes data foundations as the "hidden base of an iceberg"—without them, nothing above the surface works.

Striking the right balance

As companies deploy AI widely, they bump into an old management dilemma: how to balance strict central control with local freedom. Some bosses prefer to keep digital assets, from proprietary data to computer networks, centralised. Others give departments the autonomy to experiment and build tools tailored to their specific needs. In 2026, a middle approach is taking hold.

Figure 2.3: Neither here nor there

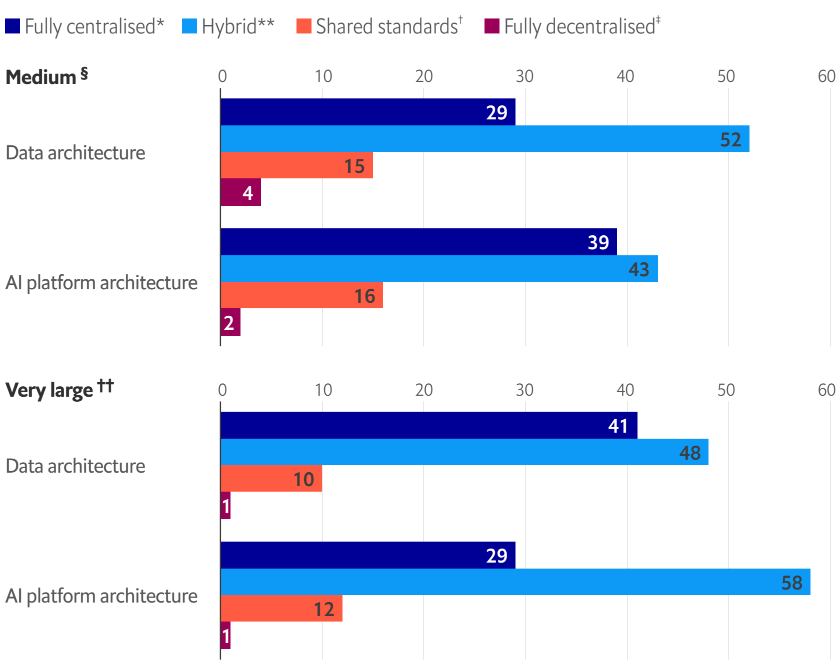
Data and AI platform architecture, % of respondents



Source: Economist Enterprise survey
* Single, organisation-wide platform/foundation with shared standards
** Central platform or foundation with some team flexibility
† Teams share a set of approved tools or systems
‡ Each team chooses and manages its own tools/systems

Figure 2.4: Size shapes the split

Data and AI platform architecture, by company size, % of respondents



Source: Economist Enterprise survey
* Single, organisation-wide platform/foundation with shared standards
** Central platform or foundation with some team flexibility
† Teams share a set of approved tools or systems
‡ Each team chooses and manages its own tools/systems
§ Global annual revenue of $1bn to less than $10bn
†† Global annual revenue of $10bn or more

Our survey shows that most firms pursue hybrid architectures, which mix tight central oversight of security and data with the freedom for local teams to develop their own AI systems. About half use a blend of centralised and decentralised systems for both data architecture and AI platforms (see Figure 2.3). Fully unified models are a distant second, favoured by about a third. Handing total control to local departments is rare: just 3% of firms fully decentralise their data, and 1.5% do so for AI tools.

The reason is often that different parts of a business have vastly different data needs. Consider Mondelēz, the American snack manufacturer. Its financial and manufacturing data is heavily regulated, requiring central systems that enforce strict standards. Marketing data, however, is messier. For example, the definitions used by each brand, from Oreo to Milka, could vary across different countries. A hybrid approach means that brands and local teams within them can manage their own data.

Our survey finds that company size is associated with where firms strike the balance. Medium-sized firms tend to impose top-down control over the technology itself, but leave their data fragmented: about four in ten unify their AI platforms, but only three in ten unify their data (see Figure 2.4). For the largest corporations, the opposite is true. Four in ten centralise data; only three in ten impose a single AI toolset.

One solution helping companies maintain order without stifling experimentation is "AI gateways"—digital checkpoints between a company's internal systems and the AI models it uses. A gateway lets managers monitor costs and enforce safety rules, while letting employees experiment freely. In its overview of technology trends for 2026, Gartner, a consultancy, predicts that by 2028, 70% of organisations deploying multiple AI models will use gateway-like platforms, up from fewer than 5% in 2024, as companies seek a way to maintain control over AI systems without blocking experimentation.[5]

5 Gartner. Top Strategic Technology Trends for 2026. 2026. Available at: https://www.gartner.com/en/articles/top-technology-trends-2026

Francis Shanahan, chief technology officer at Peloton, a high-tech wellness company, describes how his engineers built a system to orchestrate AI agents. It includes controls over both API access keys (passcodes that authenticate and control how AI connects to other systems) and the tokens AI models consume. "This provides a variety of critical governance guardrails that allow our teams to explore freely with AI while keeping the systems safe and costs down," he says.

Industry profile 2

The speed limit of data

In the car business, it is unwise to drop a powerful new engine into a rusted chassis. But carmakers and other firms in manufacturing often do exactly that with AI. Our survey finds that industrial firms are ahead in building advanced-AI systems but much less likely to have unified their data to make those systems work.

The challenge of bad plumbing is magnified by the ambition to use autonomous systems. When asked to name the biggest barrier to using AI agents, about three in ten manufacturers cite integrating AI into legacy kit. A similar share flag unreliable data from old sensors and connected equipment. About a quarter point to the challenge of managing the human workforce through an AI-enabled transition.

To bridge this gap, some firms are putting the software directly into the hands of their workers. Stellantis, a carmaker with 14 brands and a tangle of legacy systems, illustrates the tension. Annabelle Gerard, the firm's vice-president for AI and data, describes a "desynchronisation" between fast-evolving AI and data foundations that are "not yet robust". The firm's solution is to deploy a "capillary network" of more than 200 "data and AI ambassadors" inside local business units to drive execution. Kaynaz Behdin, the company's senior vice-president for digital transformation, data and AI, says the firm relies on "a clear articulation of who leads direction and who owns execution". Headquarters sets the rules and provides the technology, while local staff do the work. This autonomy speeds things up. Engineers, for instance, have cut the time needed for car-design simulations from months to a fraction of that, with some now running up to 300 times faster.

When built well, AI gateways become invisible to staff. At CRED, an Indian fintech, the gateway is wired directly into everyday software. Swamy Seetharaman, the firm's AI enabler, notes that it is not "another tool people must remember to use", but a shared layer of intelligence embedded directly into their workflow.

For Stellantis, building an AI gateway was a necessary response when a large number of the carmaker's 250,000 employees asked for access to AI models, says Kaynaz Behdin, who leads the company's digital transformation, data and AI. And at Experian, a global data and technology company, a gateway allows the firm to "apply machine learning operations and model risk management to any new AI offering, similarly to what we would do with a statistical model", notes Vijay Mehta, its executive vice-president and general manager for global solutions and analytics.

From dependency to orchestration

AI models are rapidly commoditising, driving down the price of off-the-shelf intelligence while boosting performance. Researchers at Stanford University's Institute for Human-Centered AI estimate that the cost of running a model equivalent to OpenAI's GPT-3.5 plummeted from $20 per million tokens (the fragments of text processed by models) in November 2022 to about $0.07 by October 2024.[6] This explains why in our survey executives do not worry much about paying for external AI models. Fewer than one in five consider software vendors a significant cost (see Figure 2.1).

6 Stanford Institute for Human-Centered Artificial Intelligence. AI Index 2025: State of AI in 10 Charts. April 2025. Available at: https://hai.stanford.edu/news/ai-index-2025-state-of-ai-in-10-charts

But cheap intelligence may not be as simple as it appears. Companies must weigh two looming problems: the compounding cost of high-computation tasks at scale, and the risks of relying on a single vendor.

The first has to do with the sheer financial weight of scaling AI. Although the price of basic tasks is falling, the technology is moving relentlessly towards high-computation workloads. The researchers at Stanford also find that o1, the first "reasoning" model developed by OpenAI (and now seriously outdated), was nearly six times more expensive to run than its non-reasoning equivalent at the time, GPT-4o.[7] The exact maths of these fees will constantly fluctuate, but the broader equation is clear: even as the cost per token falls, the demand for sophisticated computation rises. As companies roll out demanding features to a vast audience, the digital tolls mushroom. "When we open AI features to millions of users, API throughput and cost become real constraints," notes Youngjin Kim, chief technology officer of NOL Universe, a South Korean online travel agency.

The second problem is technological and has to do with the fierce competition among AI models. "You don't really know who the winners are going to be," says Chas Murphy, senior vice-president for direct-to-consumer data and analytics at Disney, an entertainment giant. Tying systems to a single provider is increasingly unwise, since different models excel at different tasks and, as Mr Murphy adds, "It all could change three months from now." Companies also fear being locked in because a vendor could suddenly raise prices, alter its safety rules or restrict access to corporate data.

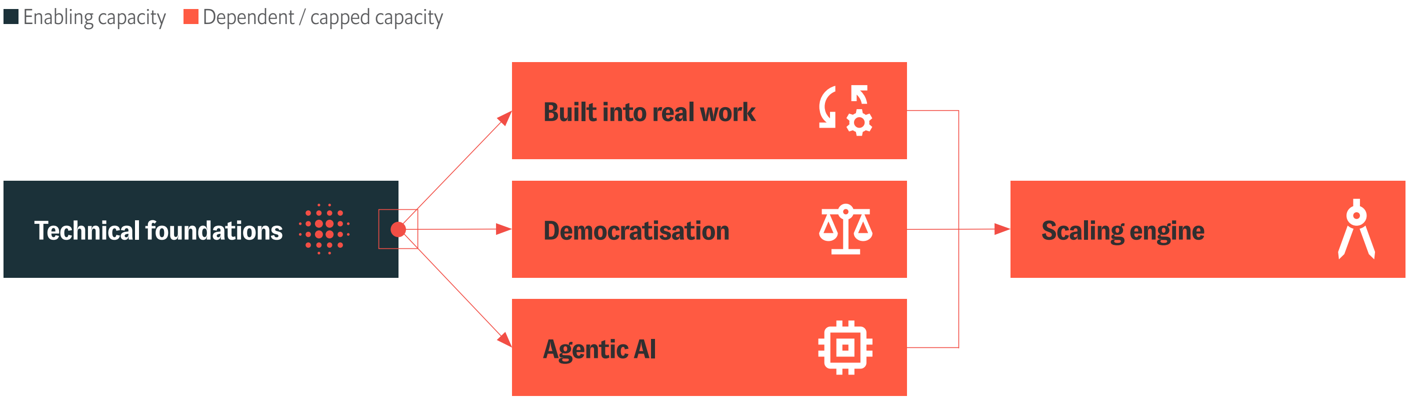
Leading firms increasingly act as orchestrators. In an industry survey by Andreessen Horowitz, a venture-capital firm, about four in ten chief information officers report running five or more AI models at once.[8] The shift towards using a diversity of AI models is also underpinned by the rise of autonomous agents, some of which tap into different models to complete complex tasks.

The broad availability of models solves the problem of dependency, but it poses a deeper commercial challenge. If competitors can buy the same algorithms, a company's commercial advantage cannot lie in models or adopting them early. To win, firms must fuse commoditised intelligence with their proprietary data and corporate know-how.

7 Stanford Institute for Human-Centered Artificial Intelligence. Technical Performance. The 2025 AI Index Report. 2025. Available at: https://hai.stanford.edu/ai-index/2025-ai-index-report/technical-performance
8 Andreessen Horowitz. How 100 Enterprise CIOs Are Building and Buying Gen AI in 2025. 2025. Available at: https://a16z.com/ai-enterprise-2025/

Dependencies and caps

Chapter 2: Technical foundations



This is why leading firms, including Disney, KONE and Atlassian, an Australian collaboration software company, are making a concerted effort to build a distinctive "control layer"—a series of guardrails and evaluation metrics alongside their data foundations. "I don't think it's going to be the next model that's the game changer," says Mr Murphy of Disney. "It's going to be how professionally they build out their guardrails and their evaluators," he argues.

Ultimately, every organisation must strike a balance between the convenience of buying off-the-shelf AI and the control of building proprietary models. "We do not believe in model monogamy," says Jose Manuel Silva of Natura. The company runs proprietary language models alongside outside services like ChatGPT and Claude, AI assistants made, respectively, by OpenAI and Anthropic, two leading AI labs, and manages the lot through an AI gateway. At the other end of the spectrum is Albertsons. Karthik Iyer, the American grocer's head of merchandising transformation and AI, says the company is building a "retail foundational model" to mirror how merchants make decisions by balancing price, promotion, assortment, space and inventory. Unlike off-the-shelf AI that solves narrow tasks, Albertsons' system integrates data, statistical learning, deep learning and AI agents to support end-to-end decisions. "Applying AI to billions of operational data points is a fundamentally different problem—both technically and operationally," notes Mr Iyer.

The foundations test

The technical foundations capacity is our benchmark's most consequential upstream dependency. Firms that exhibit strong foundations move faster on almost every other front, from scaling and democratisation to governance and agentic AI. Firms with weaker foundations find that every new AI use case they pursue lays bare weaknesses they could previously ignore.

Firms that lead share several features. They consolidate data into a form that is findable, accessible, interoperable and reusable (FAIR). They maintain data lineage so that when a model produces a bad output, the source can be traced. They build reusable integration layers, so that connecting a new AI application to existing systems does not require rebuilding the plumbing each time. And they apply the same discipline to unstructured data, from documents and recordings to emails, as they apply to structured tables.

Chapter 3: Escaping pilot purgatory

From experiment to production: the scaling engine

Corporate enthusiasm for artificial intelligence has spawned a frenzy of experimentation since the launch of ChatGPT in late 2022. Across the business world, the mandate was to move fast and explore what the technology could do.

That era is closing. Most companies now know what AI can do, but what remains unclear is how to make it work reliably across an entire business at scale, at speed and at a cost that justifies the investment. That transition is proving far harder than the experiments that preceded it. It requires solving three problems in sequence. Companies need a structured process to carry ideas from concept to production. They need the discipline to kill projects that are not delivering. And they need to embed AI so naturally in daily work that people use it without being told to. Our survey finds that most firms are still struggling with all three.

Benchmark capacities in this chapter:

Scaling engine & Built into real work

Two capacities converge in this chapter: the scaling engine and the degree to which AI is built into real work. In our framework, the scaling engine measures whether a firm has a repeatable, governed process for turning promising experiments into live systems. Built into real work measures whether internal AI systems actually fit how people do their jobs.

The two capacities work in tandem. A firm can have a mature scaling engine and still find meaningful AI use shallow if the technology sits to the side of daily work. The dependency runs in both directions: poor workflow integration caps the scaling engine's output, and a broken scaling engine prevents well-designed tools from reaching the people they were built to help.

A life-cycle approach

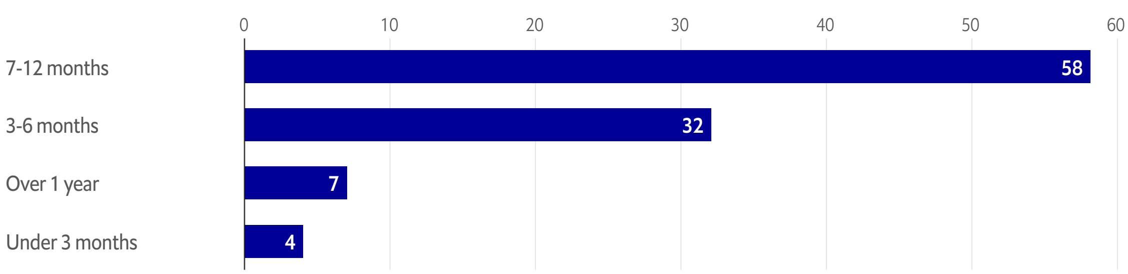
About a year ago, Ashish Agrawal, chief information officer at KONE, surveyed his company's AI efforts and was not entirely pleased. Proofs of concept had spread across the business functions unchecked. "AI here, here and there", he recalls. Adoption and its intended business benefit, he adds candidly, "wasn't happening."

One popular framework that AI boosters preach is the 30-60-90 timeline: 30 days to build a prototype, 30 to validate it and 30 to deploy it. The reality, for most large companies, falls well short. Our survey finds that about three-fifths of firms take between 7 and 12 months to move an AI project from idea to live use. Barely one in 25 manages it in under three months. About one in 14 takes more than a year (see Figure 3.1). Digital-native companies, built on software from the start, stand out. Four in ten ship within three to six months, against about a third of firms on average (see industry profile 5).

Slow deployment is partly structural. About three-fifths of firms lack a fully established AI-development life cycle—a formal process for deciding which ideas to pursue, how to test them and when to ship them (see Figure 3.2). Without one, every new project forces teams to resolve the same questions from scratch: who owns the data, how the system should be tested and who approves deployment? That rework drains time and momentum (see Capacity box 3).

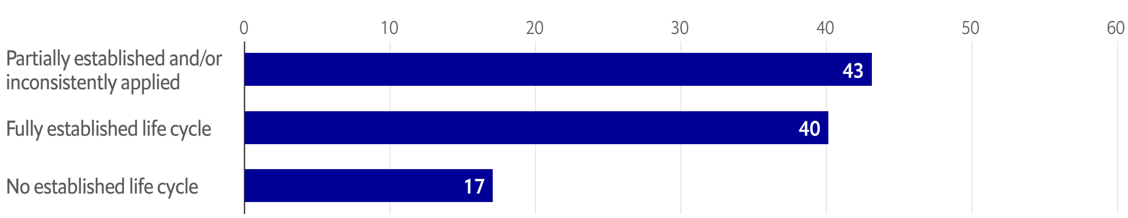
The firms furthest ahead treat strong ideas as corporate capital to be multiplied. "When we spot a solution that works, we want to reuse and scale it in different parts of the organisation," says Gabriele Ricci, chief data and technology officer at Takeda. Doing this requires co-ordination and oversight. At Mercedes-Benz, Daniel Eitler, the firm's chief AI and data officer, has set up a structured portfolio of AI initiatives reviewed by the board. "AI development is simply too fast for waterfall-style deployment," he observes.

Figure 3.1: The long road to live
Time to move an AI project from idea to live production, % of respondents



Source: Economist Enterprise survey

Figure 3.2: Life cycle, what life cycle?
Maturity of AI-development life cycle, % of respondents



Source: Economist Enterprise survey

> **"AI development is simply too fast for waterfall-style deployment."**
>
> Daniel Eitler, chief AI and data officer, Mercedes-Benz

For some companies, the pressure to build such systems is external. The most significant driver is the European Union's AI Act, which requires firms operating in the bloc to document AI systems in production, effectively mandating the kind of life-cycle discipline that many companies otherwise resist.[9]

Survival of the fittest

A good corporate practice to avoid the costs of AI experiments ballooning is to have a process that kills projects that fail and backs those that succeed. Yet having such a process remains rare. Our survey finds that firms are prone to letting their digital pipelines clog with pilots of little value. Fewer than half of organisations require their teams to link a new algorithm to a broader corporate goal. And three in five firms lack any formal process to review progress.

There is a cost to weak oversight. Shimizu Seita, chief digital officer at Tokyo Gas, recalls that investment decisions "tended to become locally optimised within individual functions", which meant that AI projects often created local efficiencies but failed to deliver broader business value. In response, the utility introduced enterprise-wide criteria and end-to-end governance to focus resources on scalable AI use cases with clear business impact.

Similarly, Natura spent nine months building sophisticated agentic AI for its human-resources operations, including multi-agent orchestration and complex workflows. The project collapsed. "We failed because we were not following the money," says Jose Manuel Silva, the firm's technology executive. "We fell in love with the architecture and lost sight of the business case." Natura now operates under a clear mandate: "between 5% and 9% of our net revenue in 2027 must be directly attributable to AI-powered models," he says.

Killing a project is harder than it sounds, because pilots that have attracted a team, consumed budget and produced a working prototype are difficult to stop. But companies must treat attrition as normal.

In some firms, the solution has been to wield the authority of chief executives to override the institutional inertia that keeps failing projects alive. At Stellantis, management cut its AI portfolio to 20 programmes, each required to show measurable value between 12 and 18 months. Kaynaz Behdin, who leads digital transformation, data and AI at the company, describes the logic: "We make a direct link between investment and value. How can we attract funds? We connect money directly to results, not to potential."

Another approach comes from Chas Murphy, data executive at Disney, who suggests that firms should approach pilots so that their failures are cheap and safe, "so that you can learn really fast".

9 EU Artificial Intelligence Act. Up-to-date developments and analyses of the EU AI Act. 2026. Available at: https://artificialintelligenceact.eu/

Narrowing the pipeline sharpens the odds for whatever survives. Suncorp started with 120 AI ideas, narrowed to 20 and cut several that failed to justify their cost. "That balance of ambition and discipline mattered," says Craig Price, who leads the firm's AI and data-science practice. One selection rule was reusability: "Instead of building everything separately, we asked: can we theme these use cases, build once and deploy many times?" Our survey finds that very large firms (having global annual revenue of over $10bn) are roughly one and a half times as likely as medium-sized ones (having global annual revenue of $1bn to $10bn) to have formal processes guiding the development pipeline, suggesting that organisational scale brings more rigorous decision-making.

Capacity box 3

Scaling engine

Strong scaling engines share three features. First, a structured life cycle: a formal process for deciding which ideas to pursue, how to test them and when to ship them. Second, disciplined attrition: the willingness of firms to kill AI projects that are not delivering. Third, design for reuse: building AI systems that can be deployed across multiple parts of the business.

Weak scaling engines exhibit the opposite. Every project relitigates the same questions—who owns the data, how will the system be tested, who approves deployment—because no agreed process exists. Projects accumulate, lose their sponsors and quietly die. Our survey finds that about three in five firms lack a fully established AI-development life cycle, and a similar share take seven months or more to move an idea into live use.

The dependency that most directly caps the scaling engine is value discipline. Without clear criteria for which projects advance and a decision-making authority genuinely willing to say no, the pipeline clogs.

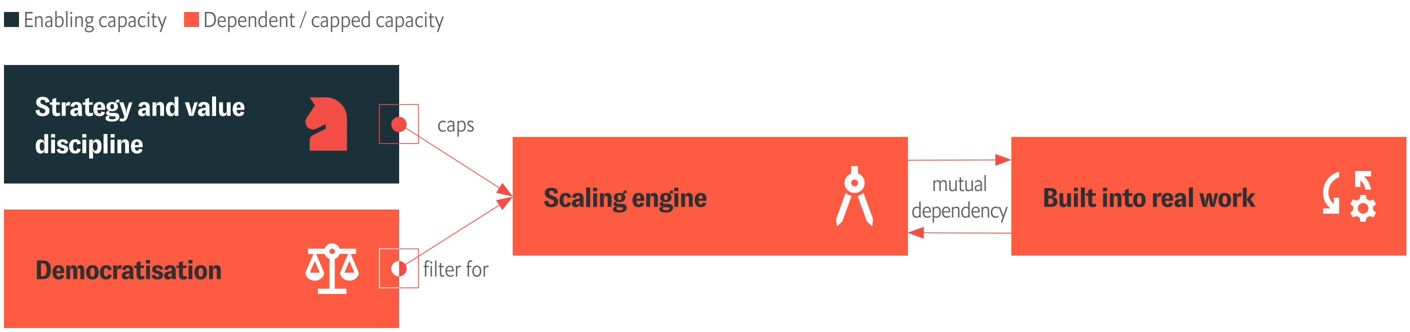
Making it stick

The surest way to waste AI is to make it harder to use than the alternative. KONE learned this quickly. Its early AI tools required field employees to open a separate application, retrieve information and then carry it manually into their existing workflow. That added effort was enough to stall meaningful use across the whole deployment, says Ashish Agrawal, the firm's chief information officer. The fix was to embed AI directly into the mobile app that field employees already used for timesheets and job reporting, so the intelligence came to them rather than the other way around. Adoption spread and complaints fell by up to 40%. "AI works best when it seamlessly integrates into the flow of any person's working day," reflects Mr Agrawal.

Albertsons frames the challenge in human terms. Karthik Iyer, the grocer's head of merchandising transformation and AI, points out that merchants must juggle prices, promotions, space and inventory when making decisions. Traditionally, they relied largely on intuition. The firm now embeds AI into their routines to piece information together and make decisions in seconds. But unless systems reflect how merchants think, "speed alone does not drive adoption," Mr Iyer says. Co-designing AI with the people who will use it narrows the gap between what technologists build and what workers adopt. At Atlassian, an Australian software company, a finance-team employee spent one Friday afternoon building an AI tool to answer colleagues' questions about travel and expense policy. "Nobody wants to answer 'what can I expense?' for the 5,000th time," observes Tal Saraf, Atlassian's chief information officer and senior vice-president for engineering. The tool worked because the person building it understood exactly where the friction lay and had access to infrastructure simple enough to remove it in a single afternoon. Through the prism of our benchmark, this illustrates how democratising AI is essential for feeding a firm's scaling engine (see Capacity box 8 in chapter five).

Dependencies and caps

Chapter 3: Scaling engine & Built into real work



Capacity box 4

Built into real work

This capacity measures whether AI sits inside core workflows or beside them. The firms with the strongest adoption share a common design principle: they bring AI to the worker. KONE embedded its technician assistant inside the mobile app engineers already used for timesheets. Atlassian built agents that operate inside Jira, a project-management tool, and Confluence, a collaboration tool. Takeda's personal assistant, used daily by seven in ten employees, is usable with no formal training.

The cap this capacity imposes is organisational. A firm can have excellent platforms, a strong scaling engine and robust governance yet see poor adoption if AI tools add friction to daily work. The dependency runs through workflow redesign. Companies must map the existing process, identify where AI fits naturally and remove any steps that might push workers back to their old ways of working. Without that investment, the most powerful AI gathers dust.

The lesson from KONE, Albertsons and Atlassian is the same: AI spreads when it fits the grain of how work already happens. But not every existing process deserves to be preserved. Mr Ricci at Takeda draws the distinction carefully. "We are entering a process economy, where the measure is how to make a process more efficient using AI." For Takeda, the inflection point came when it stopped patching broken processes with AI and began redesigning them.

Three questions that matter

The scaling engine is where the central tension of making AI work in 2026—between activity and operational strength—resolves or persists. Firms with strong scaling engines share three features: a structured life cycle, disciplined attrition and design for reuse. Without these, AI projects accumulate faster than they can be governed, evaluated or retired.

The practical test is simple. Can a firm move an idea into production within a predictable timeframe? Can it kill a project that is not delivering without a political fight? Can it take a solution that worked in one division and deploy it across a few more without rebuilding from scratch? Success is being able to answer yes to all three.

Industry profile 3

Lights, camera, administration

Artificial intelligence offers entertainers two distinct futures: eternal wealth or near-term obsolescence. To see the first scenario, consider London, where giddy crowds pay close to $2m a week to watch the septuagenarian pop group ABBA perform as virtual avatars, allowing the band to profit while putting their feet up. And to see the second scenario, consider Hollywood not too long ago when actors and writers marched in the sun, terrified that studios would scan their faces, replicate their likenesses and replace them entirely. AI can now write scripts, generate visual effects and produce polished music at a fraction of the usual cost. By hoovering up copyrighted work to automate the very soul of the business, it has turned the technology into a terrifying potential rival.

This tension is mirrored in our survey. On paper, the media and entertainment sector is well prepared to use AI, for example, by leading all other industries in having its digital systems ready to run the technology by about ten percentage points. In practice, it suffers the highest rate of staff who refuse to use advanced technologies, such as AI agents. When asked why that is, executives from the industry offer no surprises. About a third say that managing copyright, licensing and intellectual property issues in AI-generated content are a concern, and a similar share points to balancing creative control and authenticity with AI-driven automation.

The result is an industry that is embracing AI everywhere except in the thing it sells. A forecast by Deloitte, a consultancy, published in late 2025 reckons that the biggest studios would devote less than 3% of their production budgets to generative-AI tools for content creation, even as they shifted about 7% of operational spending towards AI-enabled functions such as contract management, marketing and the localisation of content for global audiences.[10] By early 2026, the pattern had hardened. In its media and entertainment outlook, Deloitte concluded that for the largest firms, the generative-AI revolution was likely to be "incremental rather than completely transformational", at least for now.[11] Yet outside the big studios the picture looks very different. Independent creators armed with cheap generative AI tools are producing near-professional-grade video at scale, competing directly for the same broad audiences as major studios. If they are not careful, Hollywood moguls might soon boast impeccably drafted contracts for films fewer want to watch.

10 Deloitte. TMT Predictions. November, 2025. Available at: https://www.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2025/tmt-predictions-introduction.html

11 Deloitte. 2026 Media and Entertainment Industry Outlook. March 2026. Available at: https://www.deloitte.com/us/en/insights/industry/technology/technology-media-telecom-outlooks/media-entertainment-industry-outlook.html

Chapter 4: Who watches the algorithm?

Governance, control and the culture of oversight

Most companies adopting AI want to talk about what the technology can do. Fewer want to talk about what happens when algorithms go wrong. Yet as firms start using AI across more and more tasks, the question of governance—who watches the technology, and how—becomes the binding constraint on how fast it can scale.

The companies pulling ahead are focusing on three main things. First, they are governing AI throughout its life cycle, from the inception of projects to well after the deployment of tools. Second, they are tailoring their rules to the danger each AI system holds, giving staff room to experiment where the stakes are low. Third, they are rethinking how accountability is exercised and what culture helps firms manage the risks AI creates.

Benchmark capacity in this chapter:

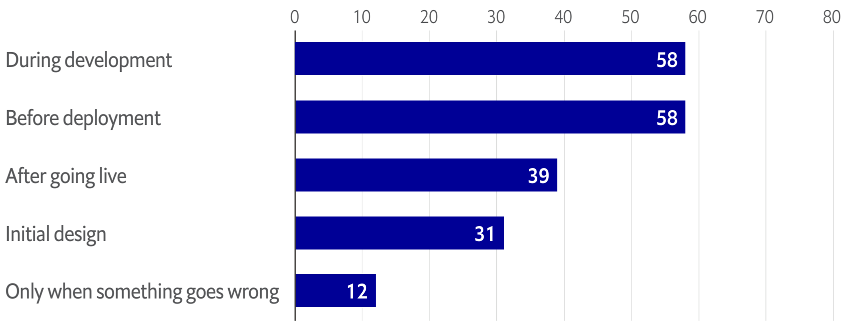
Governance and control

Governance and control is one of two foundational capacities, alongside technical foundations, whose weaknesses can cap almost every other. Unlike technical foundations, which firms tend to treat as investment decisions, governance is more often treated as a checklist.

Our survey finds that governance is strongest during the development and launch of AI systems, when it is easier to manage. But AI requires strong oversight after a system goes live as well.

Figure 4.1: The governance cliff

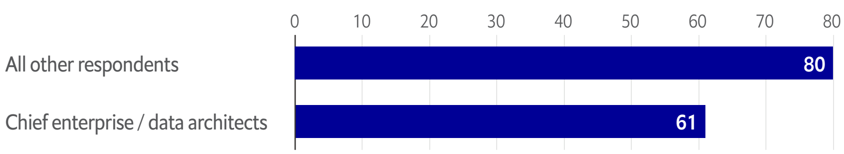
Stages at which governance reviews are conducted, % of respondents



Source: Economist Enterprise survey

Figure 4.2: The view from the engine room

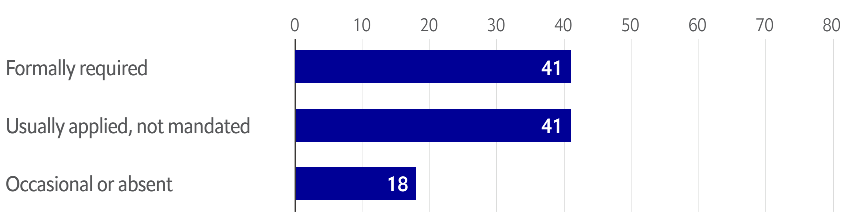
Share of executives saying their organisation's legal, compliance and governance processes support smooth AI integration, % of respondents



Source: Economist Enterprise survey

Figure 4.3: Learning nothing from the evidence

Share of firms formally requiring AI-outcome metrics to inform updates to governance policies and rules, % of respondents



Source: Economist Enterprise survey

Built for approval, not live operation

One wrong approach to governing AI is to treat it like any other enterprise software. Unlike conventional IT systems, which tended to do what they were told, AI behaves differently as conditions change. America's National Institute of Standards and Technology (NIST), which sets the country's most widely adopted AI risk-management standard, warns that deployed models can "drift" with use—quietly ceasing to match the assumptions they were built on.[12] This is why AI governance cannot stop after launch.

But our survey suggests that, for most firms, it does. Only about two in five companies have governance structures to monitor AI after going live—the very discipline NIST urges (see Figure 4.1). Fewer than a third conduct governance reviews at the initial design stage, where security can be baked in from the start. Where governance clusters is in the middle of the AI life cycle. About three in five companies review their AI systems during development and again just before deployment. One in eight admit to reviewing governance only when something goes wrong.

Chief enterprise architects—the people who actually build firms' AI systems—are the only executives in our survey for whom strengthening risk and compliance ranks as the top strategic priority. They are also the most sceptical about their organisation's readiness on every dimension we measured (see Figure 4.2). Barely three in five say their legal and compliance processes support the smooth integration of AI, compared with about four in five among other executives.

The tools to sustain robust AI governance are often scant. Only two in five firms say they operate an established AI-development life cycle that includes governance controls for regularly evaluating and rebuilding systems. And the feedback loop that ought to sharpen governance over time is largely missing: fewer than half formally require AI outcome metrics to inform updates to their rules (see Figure 4.3).

12 National Institute of Standards and Technology. Challenges to the Monitoring of Deployed AI Systems. March 2026. Available at: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-4.pdf

"We already had the discipline, the risk frameworks and the controls. We just needed to adapt them to generative AI."

Craig Price, head of artificial intelligence and data science, Suncorp

Capacity box 5

Governance and control: The life-cycle gap

A main distinction within the capacity to govern AI is whether the efforts of firms cover the full life cycle of AI projects or whether they cluster around approval gates. Our survey reveals that in 2026 most firms fall into the second camp.

Firms that lead in AI governance and control layer multiple forms of oversight, from human review by subject-matter experts to automated monitoring for shifts in model performance and even AI systems that evaluate other AI outputs. Suncorp, for example, performs monthly sampling, user feedback, back-office monitoring and drift detection.

Firms that do not perform as well see AI governance as a one-time gate. They review AI systems at inception, approve them and then rely on the assumption that they will continue to behave as designed. But AI's probabilistic nature violates this assumption. Models sometimes drift, the data on which they operate changes and edge cases multiply. A governance framework that does not follow the technology into production is a framework that governs only the version of the system that no longer exists.

Governance frameworks that do not follow through into production create risk and cap how well firms can scale AI. Weak governance is a major hindrance to deploying autonomous agents.

Companies with strong risk frameworks before the AI era are finding the transition easier. "Due to our historical investment in technical insurance pricing, we already had the discipline, the risk frameworks and the controls," says Craig Price, who leads AI and data science at Suncorp. "We just needed to adapt them to generative AI."

Failing to govern enterprise AI is already proving costly. A survey from 2025 of nearly 1,000 executives at firms with revenues over $1bn, published by EY, a consultancy, finds that virtually all companies deploying the technology had lost money to algorithmic mishaps.[13] More than three in five reported losses exceeding $1m, and the average hit was $4.4m, bringing the collective toll for the group to roughly $4.3bn. The most common culprits were broken rules, missed environmental targets and biased AI outputs. But the study also finds that taking precautions pays off. Firms with proper safeguards, such as real-time monitoring, suffered a third fewer failures.

13 EY. How Responsible AI Translates Investment into Impact. 2025. Available at: https://www.ey.com/en_uk/insights/ai/how-can-responsible-ai-bridge-the-gap-between-investment-and-impact

Matching oversight to risk

When governing AI, one size fits none. A common mistake firms make is to apply a single set of governance principles to every AI project, regardless of whether it is a chatbot summarising internal memos or an autonomous system making decisions that affect customers. Leading organisations are starting to match the intensity of oversight to the level of risk.

KONE shows how. The firm sorts AI development into three tiers, each with rising oversight. At the bottom, employees build small personal workflows with minimum essential guardrails required for data and security. In the middle, data scientists build departmental tools under IT-vetted data access. At the top, systems are fully governed by IT. Ashish Agrawal, the firm's chief information officer, kept the lower tiers deliberately light to encourage adoption. "If you want to change the culture of the organisation to be democratised, self-reliant, you need to allow certain growth," he says.

> **"You can put the process in place and get the technology right, but it won't be enough unless you get the culture right."**
>
> Chas Murphy, senior vice-president for direct-to-consumer data and analytics, Disney

Other firms arrive at the same destination by a different route. Tokyo Gas did not wait for a comprehensive policy before rolling out generative AI. The utility provided employees with approved tools and basic guardrails to encourage early experimentation and identify where the greatest value of using the technology would show up. Shimizu Seita, its chief digital officer, notes that building familiarity before "progressively strengthening governance" was "critical to achieving both speed and effective governance" as the technology's risks and benefits emerged through real usage.

For high-stakes decisions, however, strict controls are essential. Summarising a spreadsheet is a contained problem, notes Karthik Iyer, who leads merchandising transformation and AI at Albertsons. Optimising for "penny-level profit across billions of data points is a different scale altogether," he says, and "a single mistake can wreck profits." This is why Albertsons' merchants review AI recommendations and enforce pricing and promotional guardrails that meet the grocer's financial, commercial and legal boundaries. "Governance is not about slowing things down," Mr Iyer says. "It is what makes this level of speed and scale viable in the first place."

The culture test

Vigilant staff can provide an ongoing check on AI. Disney learned this in practice. Three days after one of its retrieval-augmented generation systems (of the sort that combine a language model with access to internal data) had cleared every approval gate, from design review to validation and launch, a real-life test returned a confident hallucination. "You can put the process in place and get the technology right," reflects Chas Murphy, data executive at Disney, "but it won't be enough unless you get the culture right." Formal checks mostly catch the risks firms can foresee. The rest depends on whether staff stay alert once an AI system is live.

There is a parallel with past corporate risks. Firms did not slash financial fraud by writing stricter rules alone. What made a significant difference was teaching people to spot trouble early, making it easy and safe to raise concerns and stitching checks into the rhythm of daily work. AI calls for a similar approach. Many of its failures tend to surface not in the model itself but at the point where it meets real work, resulting in a biased shortlist, a hallucinated clause in a contract or a customer-service answer that is fluent and wrong. Only a workforce trained to question AI, and that feels safe doing so, can anticipate such collisions.

Few manage vigilance well. Our survey finds that about four in five executives say their firm has organisation-wide policies governing AI use. But practice is far messier. A study from late 2025 sponsored by IBM, a software giant, finds that while four in five American office workers use AI on the job, only about one in five stick exclusively to tools their employers have approved.[14] This gap between policy on paper and AI use in practice has a name: shadow AI.

Workers who hide their AI helpers are usually acting on rational self-interest. If admitting to a clever shortcut could mean a heavier workload, a sceptical manager or a redundancy, why volunteer? The result is that experimentation with AI goes underground, where the firm can neither learn from it nor govern it.

Maria Macuare, chief data officer at Mondelēz, is blunt: "In the industry we live under the illusion that we can limit exposure. But I think we could be deceiving ourselves as to how much we're really controlling." The problem is compounded by speed. Tal Saraf, chief information officer and senior vice-president for engineering at Atlassian, captures the treadmill: "The thing that worries me is literally every single day a new thing comes out, and people want to play with it."

A smarter response works on three fronts. The first is incentives. Employees must know that using AI meaningfully will advance their careers, not merely pile on extra work or make their roles dispensable. In our survey, around four in five executives say that incentives in their company reward successful AI innovation. One illustration of how to set the right incentives comes from CRED, an Indian fintech firm. Swamy Seetharaman, who works on enabling AI at the firm, explains how as the technology absorbs repetitive tasks, people move "from producing to deciding, from doing to exercising judgment, context and taste".

The second is psychological safety. Staff who question an algorithm, report a mistake or escalate a concern should be prized as the sceptics who stop a firm from blindly trusting AI. Wendy Batchelder, chief data officer at Centene, is explicit: "We need contra perspectives to help us ensure we are considering all angles."

Capacity box 6

Governance and control: When workers go underground

Shadow AI, the unsanctioned use of tools the employer has not approved, is one clear symptom of a governance culture that fails to earn trust. If a firm's workers hide their AI helpers, it is unlikely that leaders have provided the right incentives and psychological safety to empower meaningful and safe AI use.

Firms with a strong culture around AI make it clear that using the technology to become more productive will lead to more interesting work, not a redundancy notice. They ensure that leaders visibly use the tools themselves. And they champion staff who question AI outputs.

There is a dependency between governance and work redesign and skills. Firms that redesign roles around AI are likely to find that shadow AI use diminishes. Firms that deploy powerful tools without adjusting the incentive structure find that governance becomes a policing exercise that is hard to win.

14 IBM. Is Rising AI Adoption Across the US Workforce Creating Shadow AI Risks? November 2025. Available at: https://www.ibm.com/think/insights/rising-ai-adoption-creating-shadow-risks

The third is leadership by example. When bosses visibly experiment with AI, they signal both permission and expectation. And that, according to several of the executives we interviewed, is the single biggest lever for cultural change. At Nasdaq, AI features on every weekly leadership agenda. The chief executive has insisted that he, too, must become "AI savvy". Mr Seetharaman of CRED goes further: "When senior leaders lead from the front as builders, using AI for planning, reviews, synthesis and creation, AI shifts from a nice-to-have capability to an operating hygiene expectation."

Who holds the reins

Well-designed rules and a vigilant workforce still need a clear answer to a blunt question: when an AI system misfires, whose phone rings? As companies scale the technology, many are discovering that the old organisational chart does not provide one. Three things help.

The first is structure. Central control keeps the rules consistent but clogs the pipes. "Our centralised centre of excellence worked, but it created an 'us versus them' dynamic," says Jose Manuel Silva, head of technology at Natura. He moved to a hub-and-spoke model that pushes some authority out to local teams while the hub retains governance and security borders. Mr Silva is candid about the trade-off: "Without proper governance—clear boundaries around data access, model auditability and ethical guardrails—decentralisation becomes dangerous. The art is distributing speed without distributing chaos."

One practical tool for managing that tension is sandboxes—walled environments where staff experiment with new AI tools. At Nasdaq, "we create a sandbox where there is no firm data," says Angie Ruan, the firm's divisional technology chief.

The second is seniority. Mercedes-Benz merged its data and AI leadership into a single role—chief AI and data officer—after recognising that the two functions could no longer be run apart. "One without the other will not work any more," says Daniel Eitler, who holds the role. Responsible AI is woven into project management. "It's not seen as a blocker," says Mr Eitler. "It's part of the process—not just a checklist or an Excel spreadsheet to be filled in and passed on."

Nasdaq adds a financial gate to the logic of having data, analytics and AI leadership under one executive. "We have an ROIC framework—return on invested capital," says Ms Ruan. "We plug that into the final governance process before anything actually scales."

The third is teamwork. At Atlassian, oversight spans legal, privacy, security, trust, product and engineering—what Mr Saraf calls "a team sport". Takeda goes further still. "The scope of the governance body is not just enforcing policies," says Gabriele Ricci, the company's chief data and technology officer. "It's about creating the right maturity for the organisation to adopt AI at scale."

Sowmya Gottipati, corporate chief information officer at Estée Lauder, a beauty conglomerate, draws a distinction that several executives echoed. "The governance and the framework—that's under my control, within the technology team," she says. "The change-management piece is with the entire business."

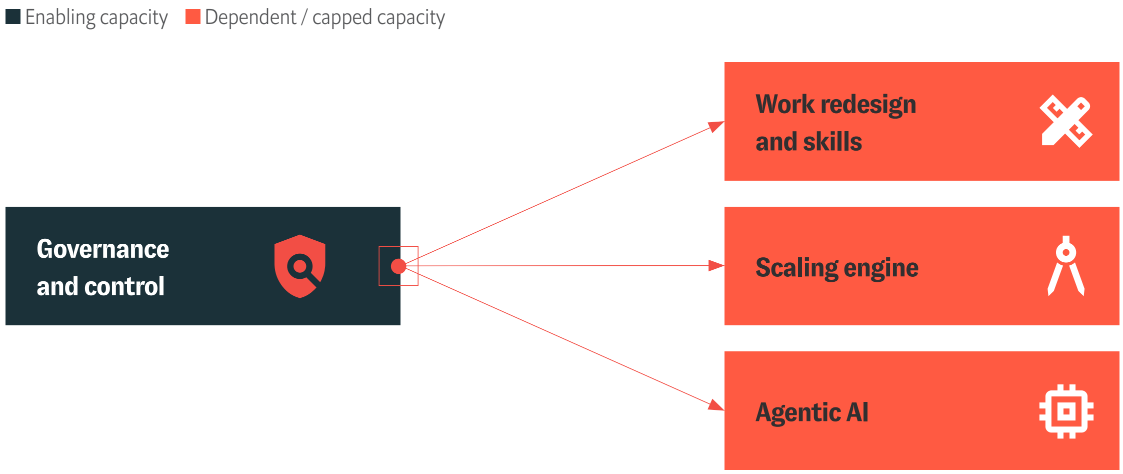
None of this works without one final ingredient: a kill switch with a name on it. At CRED, every AI tool requires a named business owner and humans with explicit authority to pull the plug. The value of such clarity shows when things go wrong. When a deployment caused an incident at Atlassian, it took 14 minutes to fix: a human spotted the need to roll back, and automated agents carried out the order.

Regulation is poised to make the governance agenda more demanding. In a report from October 2025, the Financial Stability Board, an international body, urged financial authorities worldwide to develop more robust monitoring of AI adoption and its associated vulnerabilities, warning that third-party dependencies and model-governance challenges pose systemic risks.[15]

15 Financial Stability Board. Monitoring Adoption of Artificial Intelligence and Related Vulnerabilities in the Financial Sector. October 2025. Available at: https://www.fsb.org/2025/10/monitoring-adoption-of-artificial-intelligence-and-related-vulnerabilities-in-the-financial-sector/

Dependencies and caps

Chapter 4: Governance and control



Industry profile 4

The extra hurdle

Banks and insurers share the corporate world's struggles with AI, from rusty digital pipes to stubborn workplace cultures. Yet even when they solve those problems, financial firms face an extra hurdle: the watchful eye of the regulator, and the unsettling fact that nobody yet knows what will make it blink. In our survey, more than a quarter of financial executives name regulation as their biggest barrier to autonomy, above any other obstacle. This added caution is expensive. Nearly three in five financial firms cite human review and compliance as a top ongoing cost, compared with half across all industries.

Financial-services companies have the highest share of firms reporting lagging return on investment—about one in 12, compared with one in 25 across industries. Yet this sluggishness masks a clear sense of what the technology can do for financial firms. Nearly half say they want to hunt better for fraud with AI and more than two in five point to automatically settling insurance claims.

Vijay Mehta, global solutions and analytics executive at Experian, captures the mood. Because there are no clear rules yet on machine-made decisions, humans remain firmly in control of the firm's regulated work. But behind the scenes, Experian is building the digital plumbing needed to verify and control software programs. Mr Mehta describes "know your AI", borrowing the shorthand from the banking world's "know your customer" rules.

American Express, a payments company, took a deliberate sequence. "We moved slowly to start with, to make sure that we built all of the right controls—people, process and technology—along the way," says Jason Sharples, senior vice-president and chief information officer for global merchant and network services at the company. With those foundations in place, the focus has shifted to change management, he adds.

Why governance pays

Governance and control is the capacity most likely to be underinvested in and most costly when it fails. The firms with the strongest governance treat it as an operating discipline. They layer automated monitoring, human review and cultural incentives into a system that evolves as the technology does. They assign clear accountability to people with the authority to act. And they treat staff as a distributed sensor network whose scepticism and alertness add value. The pay-off is not only risk reduction but also more confident experimentation and scaling.

Chapter 5: Rewiring the firm

Work redesign, skills and the operating model

The rise of AI has sparked a fascination with the history of corporate invention. Economists go back to the industrial past to see how earlier general-purpose technologies reshaped work and the economy. One common destination is 19th-century factories, where the productivity gains of electrification stagnated for decades until managers completely redesigned floor layouts, processes and organisational structures. Case studies of more recent times include one by researchers at the Massachusetts Institute of Technology (MIT) who in 2015 examined football-makers in Pakistan.[16] The researchers found that workers stubbornly resisted a new wastage-reducing technology because it disrupted their familiar routines and piece-rate wages.

These echoes of the distant and more recent past reveal plenty about AI and the future of white-collar work. Unlocking the full economic potential of AI demands the gruelling task of rewiring corporate culture, workflows and human incentives. This requires time and money, making balance sheets uncomfortable. It also involves deconstructing job descriptions and rethinking how expertise flows through the organisation. And it raises a structural question about how AI work is organised and supported across a firm.

Capacities in this chapter:

Work redesign and skills, Democratisation & Operating model

Three capacities underpin this chapter: work redesign and skills, democratisation and the operating model. The first is work redesign and skills, and it reflects whether firms have remade roles, tasks and incentives around AI, including the difficult work of separating what algorithms can perform from the tasks that require human judgment. The second, democratisation, is about whether non-technical staff can access and use AI safely, with the right training and infrastructure. The third is about a firm's AI operating model and ecosystem. It includes questions about who decides, who builds and who runs AI across the organisation as well as how that work is funded, and whether the structure rewards reliable scaling or tolerates fragmentation.

In our framework, these three capacities sit at the centre of AI operationalisation, as weak performance in any of them caps what AI systems can deliver in practice.

16 David Atkin, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen. Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan. July 2015. Available at: https://dspace.mit.edu/bitstream/handle/1721.1/113695/On%20the%20origins.pdf?sequence=1&isAllowed=y

The weight of co-invention

Artificial intelligence is a technology that makes a bad first impression on most chief financial officers: it promises to make their numbers fly yet often first sinks them. Before firms see productivity gains from using AI, they typically see the reverse: costs rise, workflows are disrupted and employees spend time learning new ways of working. This forms the productivity J-curve, a dip in measured output before the eventual surge. It is the unavoidable price of what Erik Brynjolfsson, a leading economist of the digital age, calls corporate co-invention—the complementary investments in people, processes and structures that make AI pay off.

Our survey captures some of this ambition. About three-quarters of respondents say their organisation has redesigned work processes to suit AI. A similar share has redesigned job descriptions (see next section and Figure 5.1). But the numbers are inconsistent with spending. Most point to things like data infrastructure, human oversight and quality assurance as their heaviest ongoing costs, with only 4% citing employee upskilling (see Figure 2.1 in chapter two for the full picture). In benchmark terms, this gap signals that work redesign is claimed far more often than funded (see Capacity box 9).

The executives interviewed for this research describe the J-curve in vivid terms. Ashish Agrawal, chief information officer at KONE, captures the tension between investment and evidence. "We all face a duality in terms of investing in AI and the return it brings to the company's profitability or growth," he says, adding that "it's always a derived benefit." Maria Macuare, chief data officer at Mondelēz, puts it more bluntly. "If anybody is saying that they are really happy with their ROI, they are probably lying," she says. "Behind closed doors, most managers are frustrated with how quickly AI is adopted and how much money is being left on the table."

Our survey suggests that firms are mindful of the usual productivity J-curve. When asked about their top goals for AI over the next two years, executives are more likely to prioritise AI-enabled innovation and transformation than plain cost-cutting.

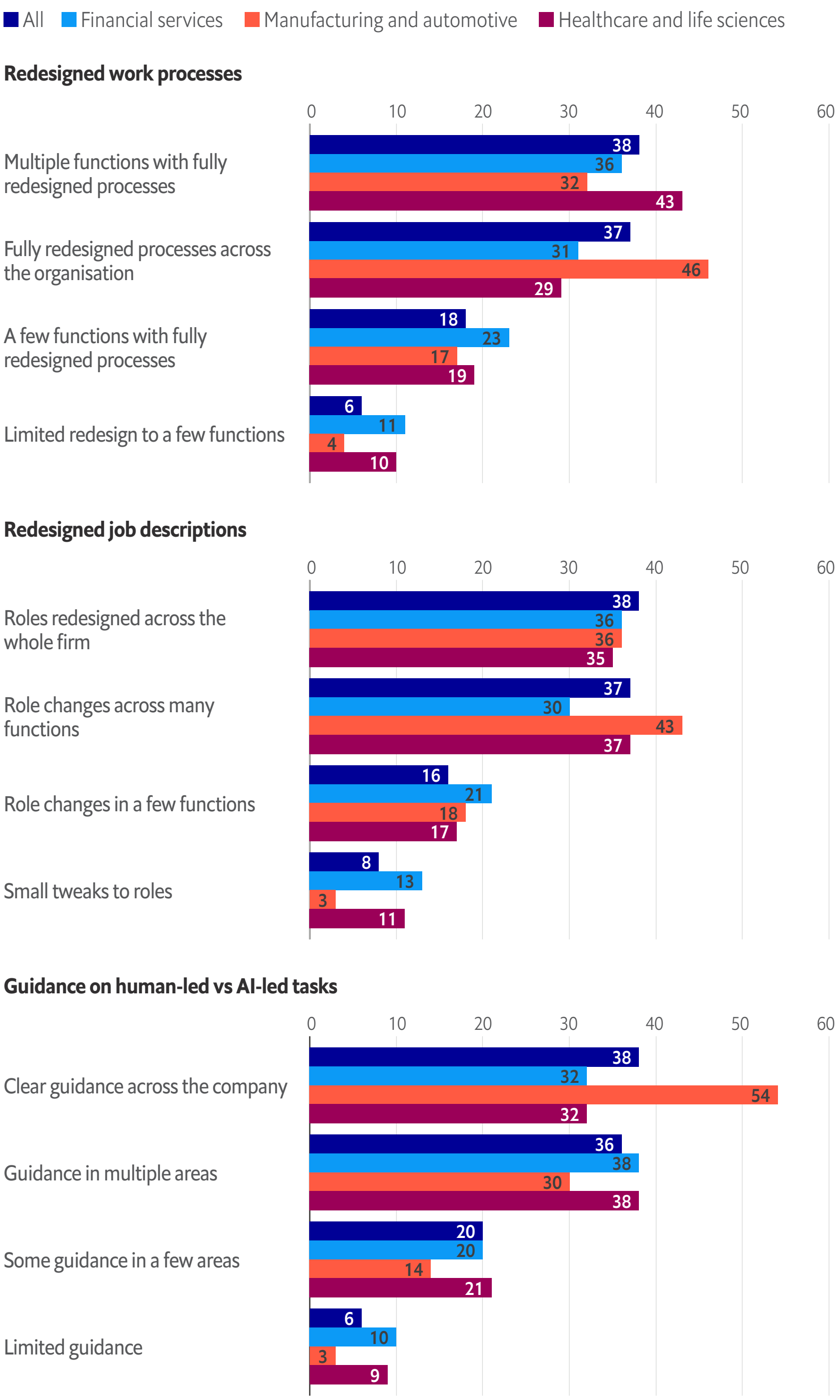
We find that most firms believe that they are starting to escape the dip and are broadly optimistic—perhaps dangerously so. About four in five executives say their AI adoption is ahead of plan, and a similar share report that return on investment is tracking ahead of expectations. A survey by McKinsey, a consultancy, from 2025 adds context.[17] Among nearly 2,000 firms across 105 countries, only 6% say AI accounts for at least 5% of their earnings and has created significant overall value. (Two in five report any enterprise-level impact on earnings at all.) The gap between perceived progress and measured impact is wide.

The firms pulling ahead started their data and cloud journeys years before advanced AI arrived. Craig Price, who leads the AI practice at Suncorp, traces his firm's readiness back more than a decade. The company consolidated its systems onto a single platform, unified its data estate and built an experienced data-science team. "What all that meant was that we started generative AI prior to ChatGPT coming out, and could move comparatively quickly to leverage this technology when it was clear the value was there," he says. The lesson, through the lens of our benchmark, is that strong technical foundations (see Capacity box 2) lower the cost of co-invention by removing the plumbing problems that would otherwise be costly.

17 McKinsey. The State of AI in 2025: Agents, Innovation, and Transformation. November 2025. Available at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Figure 5.1: Redesign on paper, gaps in practice

Share of firms reporting organisation-wide activity on AI-related work redesign, % of respondents



Source: Economist Enterprise survey

Breaking work into its atomic units

Firms that extract real value from AI tend to deconstruct jobs into their smallest component: the task. Every occupation comprises dozens of discrete activities. Mapping them allows leaders to identify precisely which tasks should stay with humans, which can be fully automated and which work best as a collaboration between the two. This "task-based" framework, developed by labour economists including David Autor and Daron Acemoglu of the Massachusetts Institute of Technology, has become the standard lens for technological change.

Our survey shows that the rethinking of jobs is well under way. About three-quarters say they have conducted broad or organisation-wide activity to redesign work processes to make the best use of AI's strengths, and a similar share say they provide clear guidance on which tasks should remain human-led versus AI-led (see Figure 5.1). Three-quarters also report redesigning job descriptions, but the shift is easier to see in some sectors than in others. In manufacturing and automotive more than half of respondents say guidance on human-led versus AI-led tasks is already organisation-wide, compared with about a third in healthcare and life sciences. The gap makes sense. Factory work is easier to instrument, observe and standardise. Clinical, legal and research work contains more tacit judgment, more edge cases and higher penalties for error.

Daniel Eitler, chief AI and data officer at Mercedes-Benz, describes this granular work as central to the carmaker's strategy. "We are working very intensively with HR on job-profile redesign," he says. Mercedes-Benz has also introduced "AI champions" who act as translators between technical teams and the wider workforce, identifying which tasks in their area are ripe for rebuilding.

Across the road at Stellantis, a new balance between human and AI emerges. Kaynaz Behdin, who leads transformation, data and AI at the company, says that "while AI can generate code and even design a car from a prompt, it does not replace human judgment." In car design, for example, "prompting can generate endless shapes—but proportion, emotion and brand identity still come from human designers," Ms Behdin notes.

Brian Bischoff, chief technology officer at CapTech Ventures, a technology consultancy, describes a technique that directly addresses the fear such redesigns provoke. He calls it "job crafting"—employing the very people whose roles are being reshaped to help redefine what those roles look like. "If you do that early on, the adoption rates are much higher," he says, "because employees feel like they're part of the process, and they understand what the end-game is."

The task-level approach also prevents the common trap of bolting AI onto a role and hoping for the best. Mr Brynjolfsson calls it the Turing Trap and argues that the drive to make AI indistinguishable from human workers leads firms down a dead end.[18] Systems designed to replace human labour rather than augment it tend to yield diminishing returns on cost-cutting, while destroying the complementary human judgment that makes AI systems improve over time. The more productive goal is augmentation: building "centaur" teams where humans and machines each contribute what they do best.

Several executives we spoke to say their firms build centaurs in practice. Wendy Batchelder, chief data officer at Centene, describes a pilot for the company's care managers. Care managers engage directly with members and spend a considerable amount of time navigating different systems to prepare for meetings with members. Using the data available to the care manager, the AI pilot distils that information into a pre-made agenda, flagging the five most urgent issues in order of priority. "Early reports are that it is saving them up to 50% of the administrative portion of their day," she says. At Experian, agile software teams have gone further still, formally assigning AI agents story points and requirements alongside human developers. "We have teams that include agents doing specific tasks around quality assurance, testing and documentation," says Vijay Mehta, the company's general manager for global solutions and analytics.

Forming centaur teams also makes AI's contribution more easily measurable. Because firms can compare a human-AI pair against a human working alone, they can isolate what the technology actually adds—and spot where it subtracts. The most rigorous firms are building the financial architecture to support exactly that. Gabriele Ricci, chief data and technology officer at Takeda, describes moving towards what he calls a "process economy", assigning every workflow its own cost structure that explicitly accounts for both human and AI labour. "You need dedicated teams for process improvement and full telemetry to understand and optimise costs," he says. "It is a major shift—less about technology than about the organisation and processes underneath it."

A parallel concern is cognitive decline—the risk that junior staff, raised on AI-assisted tools, never develop the foundational skills that senior colleagues rely on. One emerging response is structured pair programming—pairing a junior and senior engineer so that AI accelerates output without eroding the mental models beneath it.

18 Erik Brynjolfsson. The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence. Spring 2022. Available at: https://www.amacad.org/sites/default/files/publication/downloads/Daedalus_Sp22_19_Brynjolfsson.pdf

Capacity box 7

Work redesign and skills

The capacity to redesign work and skills most directly shapes whether AI's J-curve bends upward or flattens into a plateau. Firms should set clear task boundaries and accountability, provide meaningful training and find ways to show measurable gains. The reality, in most firms, falls short.

Three types of companies emerge. The first group has decomposed jobs into tasks and explicitly assigned each task to a human, a machine or a collaboration between the two. Centene's care-management pilot, Experian's agile teams with AI agents assigned story points and Takeda's "process economy" with dedicated cost structures for each workflow all represent this approach. These firms can measure what AI adds because they have isolated its contribution.

The second group has redesigned job descriptions on paper but not in practice. Our survey suggests this group is large. About three-quarters of respondents report redesigning job descriptions, yet only 4% name employee upskilling as a significant ongoing cost. It could be that firms find upskilling cheap, but the mismatch is more likely to suggest that many "redesigns" have not been accompanied by the level of investment in training, incentives and workflow change that would make them meaningful.

The third group has not yet started. These firms deploy AI tools and hope that workers will figure out how to use them. The evidence from our interviews is that this approach produces uneven adoption, with enthusiasts racing ahead and sceptics falling further behind.

Levelling the playing field

Redesigning jobs at the level of tasks is one half of making AI work for people. The other half is about broadening who can use it. AI tends to lift the floor of capability across an organisation, compressing the gap between novice and expert. That is the democratisation promise of the technology. But it also raises the ceiling for those with the judgment to use it well and, counter-intuitively, makes their role more critical.

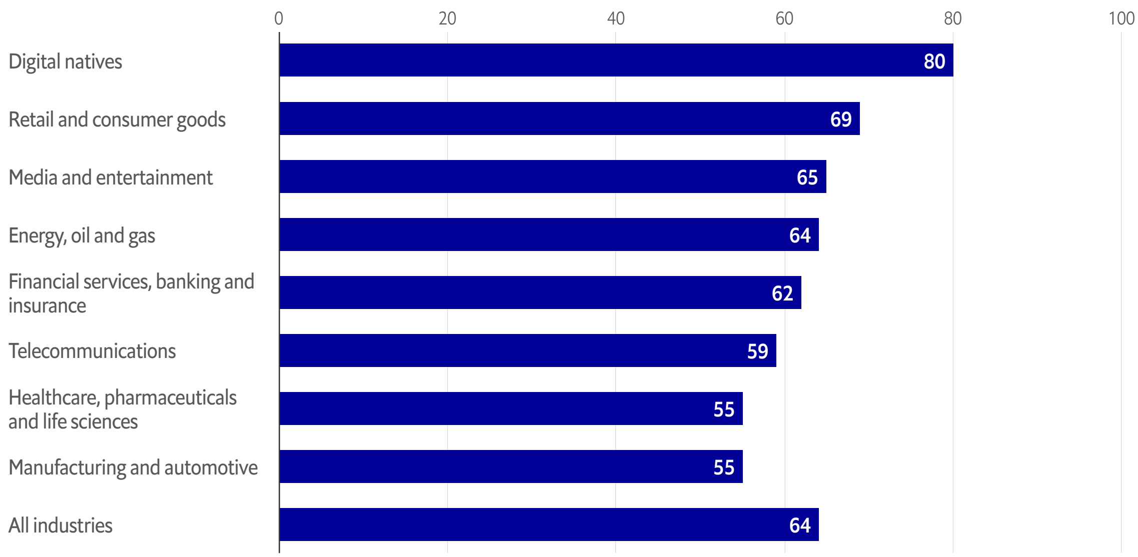
A run of studies provides early evidence. One, from 2024, by the Sloan School of Management at the Massachusetts Institute of Technology, tracks software developers at Microsoft, Accenture and a Fortune 100 electronics-maker as they gain access to AI coding assistants.[19] Overall output rose by 26%, but junior and recently hired developers saw gains of between 27% and 39%; senior developers gained only 8% to 13%. "For those who are more experienced we actually don't see much of an effect," said Mert Demirer, one of the authors. Another study, by Mr Brynjolfsson and his colleagues, tracks 5,172 customer-support agents through the introduction of an AI assistant.[20] They find that it raises productivity by 15% on average, but the gains fall almost entirely on the least experienced workers, who improve both speed and quality, while the most skilled see small speed gains and even small declines in quality.

19 MIT Sloan. How Generative AI Affects Highly Skilled Workers. November 2024. Available at: https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-affects-highly-skilled-workers

20 Erik Brynjolfsson, Danielle Li, and Lindsey Raymond. Generative AI at Work. May 2025. Available at: https://academic.oup.com/qje/article/140/2/889/7990658

Figure 5.2: Closing the data divide

Share of non-technical employees with active self-service access to data or AI-assisted tools, % of respondents



Source: Economist Enterprise survey

In effect, AI performs skill compression. It encodes the tacit knowledge of top performers in a company and makes it accessible to those who would otherwise take years to acquire it. And in narrowing experience gaps, AI looks less like a productivity tool and more like an equaliser.

This equalising force shows up most tangibly in how firms handle data and tools. In our survey, about two-thirds of firms across industries say that non-technical staff have active self-service access to data, either through AI-assisted tools or automatically generated recommendations (see Figure 5.2). Among digital-native firms—those born on software and fluent in data—that share rises to four-fifths. And in healthcare, often weighed down by legacy systems, fragmented records and tight regulation, the share sits at just over half.

The gap reflects how much democratisation depends on what came before it. Firms that invested early in clean, accessible data infrastructure find that AI can now put meaningful insights directly in the hands of a salesperson, a care manager or a logistics planner. This is one of our benchmark framework's clearest dependencies: democratisation is capped by the strength of a firm's technical foundations (see Capacity box 2), because staff cannot use data they cannot find, and they cannot trust tools built on data they know to be patchy.

Access without structure, however, creates its own problems. One wrong conclusion from the finding that AI benefits the least experienced most is that this makes senior experts dispensable. Quite the opposite. Senior workers remain essential precisely because they supply what AI cannot yet reliably produce: the high-quality judgments that train AI systems, the standards those systems are measured against, and the ability to catch failures before they become costly.

Capacity box 8

Democratisation

The companies strongest in democratising AI combine low barriers to access with clear boundaries on use. KONE's three-tier model, which includes personal tools with minimum essential guardrails for data and security at the bottom, IT-vetted departmental tools in the middle and fully governed enterprise systems at the top, illustrates this. The firm now has 4,000 citizen developers building their own tools, with cyber-security and data privacy as the only hard limits throughout.

Weaker performers in this capacity provide access without structure. These firms roll out tools broadly, celebrate high adoption numbers but discover, often too late, that dozens of teams have built overlapping agents, consumed unexpected amounts of compute or exposed sensitive data to unsanctioned models.

The dependency of this capacity runs in two directions. Democratisation hinges on strong technical foundations (staff cannot use data that is hard to find) and on governance (guardrails must exist before access is widened). But it also feeds back into the scaling engine. The more people who can safely experiment, the richer the pipeline of ideas that the firm can evaluate, theme and scale.

Kim Hales, chief data and AI officer at NRG Energy, one of America's largest independent power producers, is trying a structural answer. His firm has created a role it calls the "business developer", a non-technical employee paired with a senior engineer in the software-development life cycle. "After someone totally non-technical has spent some time with our senior engineers, they are able to deliver a dramatic step-change in the speed of deployment. Things that used to take months can now just take weeks," he says. The approach compresses the gap between domain expertise and technical fluency without requiring either side to become the other.

KONE's technician AI assistant illustrates how the learning curve is flattening. Built on Anthropic's Claude large language model, it serves more than 8,000 field technicians across over 40 countries, distilling three decades of engineering knowledge into a tool anyone can consult. Customer complaints have fallen by up to 40% in some markets. "The AI is the buddy to the technician—not the replacement," says Mr Agrawal.

Another consideration is not to leave any group behind. Ms Batchelder at Centene raises a concern that our survey does not capture: women are adopting AI tools more slowly than men. "Do we have a microcosm of that, and are there different approaches we need to take from a training and enablement perspective?" Her concern has empirical backing. An analysis from 2024 by the Oliver Wyman Forum, a think-tank linked to Oliver Wyman, a consultancy, found that men were about 10% more likely than women to use generative-AI tools regularly.[21] The gap was widest in older age groups and in industries where women held fewer technical roles.

21 Oliver Wyman Forum. Women Are Falling Behind On Generative AI In The Workplace. Here's How To Change That. April 2024. Available at: https://www.oliverwymanforum.com/artificial-intelligence/2024/apr/women-are-falling-behind-on-generative-ai-in-the-workplace--here.html

Who decides, who builds and who runs

Every firm deploying AI builds two architectures, even if it only consciously designs one. The first is the technical framework of data platforms and hybrid infrastructures (see chapter two). Its shadow, and inevitable twin, is the organisational structure that governs it. That second architecture is about who decides which AI projects to pursue, who builds them, who runs them in production and how the money flows.

The most common structure in 2026 is a hub-and-spoke model. In it, a central team sets standards, maintains shared platforms and enforces governance, while distributed teams in business units build and run the applications closest to their work. At Natura, Jose Manuel Silva moved from a centralised centre of excellence to a federated structure after the old model created what he calls an "us-versus-them dynamic". The hub now defines guardrails while the spokes own applications. "That decision unleashed more speed," Mr Silva says, though he adds a caveat: "Without proper governance, it's a bit dangerous."

Mercedes-Benz runs a variant. Mr Eitler maintains a network of "AI officers" embedded in each business unit. They report locally but align with Mr Eitler's central strategy. The arrangement ensures that AI projects are shaped by people who understand their daily business—"For instance, I personally have never built or sold a car," Mr Eitler notes—while the centre prevents duplication and keeps a board-level view.

Takeda adds a further layer. Its "AI forward" programme connects distributed teams through a shared governance body, a community platform used by 4,000 employees, and an "agentic academy" that trains both business-process owners and engineers. Annabelle Gerard, who oversees AI and data at Stellantis, faces a similar challenge at greater scale: "AI is often considered IT, and that means centralisation. But AI sits close to business domains, so our mission is to explain how important it is to transition to a federated approach."

Our survey offers a quantitative glimpse. About half of respondents use a blend of centralised and decentralised systems for their AI platforms. Very large firms are roughly one and a half times as likely as medium-sized ones to have formal processes governing the development pipeline, suggesting that scale eventually forces the operating-model question into the open. But size alone does not settle it. What matters is whether the answer to "who owns this?" is explicit, stable and understood across the business.

Culture eats strategy

Platforms and training programmes matter. But the executives we spoke to, almost without exception, return to the same point: culture determines whether AI scales or stalls. Our interviews and survey point to three broad groups in any organisation: enthusiasts ready to leap, sceptics paralysed by fear and fence-sitters who have not yet seen the value. Guardrails and governance help contain the first group and reassure the second. But the third, often largest, group needs proof. "It's evangelism," says Chas Murphy, data executive at Disney. "People just like you are using AI this way, and look at what it's unlocked for them." In our survey, about four in five executives believe that staff are open to change and experimentation with AI.

Capacity box 9

Operating model

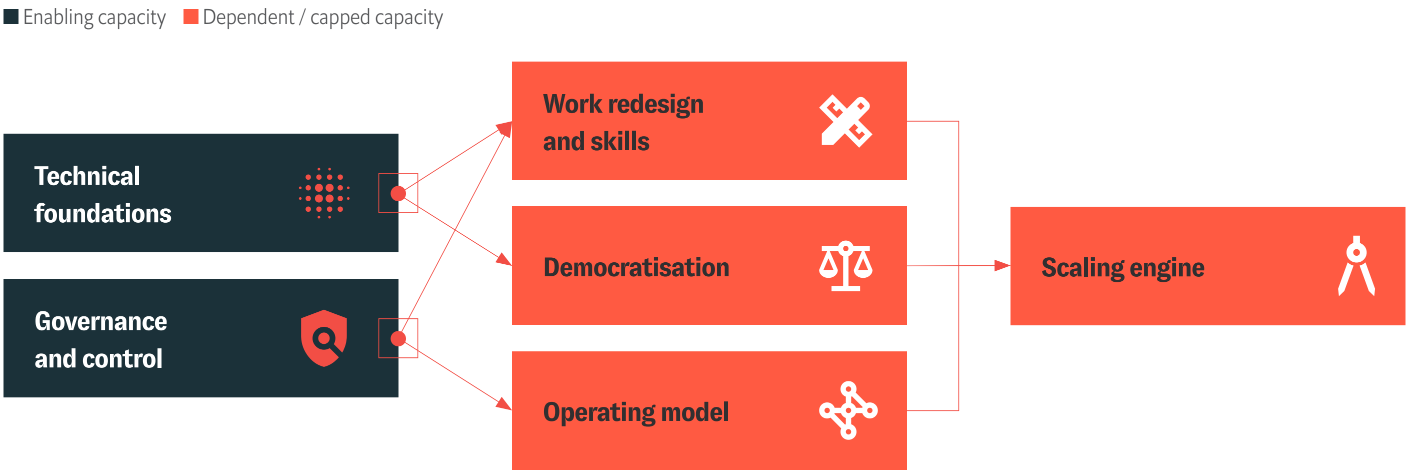
The operating model measures how AI work is organised, funded and governed across teams and suppliers. Its yardsticks are plain: the answers to who decides, who builds and who runs are explicit; product and platform teams are not assembled ad hoc for each project; and incentives reward reliable scaling.

Three broad approaches emerge. The first is a hub-and-spoke model in which a central team sets standards while business units own applications. Its strength is that it offers speed with coherence, but there is a risk that the spokes drift from the hub's standards over time.

The second is full centralisation, where a single AI team controls all development. This produces consistency across the organisation but disconnects AI from the business problems it is meant to solve, risking less meaningful use of the technology. The third approach is ungoverned decentralisation, where business units build independently without shared standards. This produces speed in isolated pockets but makes reuse, governance and cost control hard.

Dependencies and caps

Chapter 5: Work redesign, democratisation & operating model



There are a few ways to shape corporate culture. Consider Takeda. Four years ago, the company added a line to its corporate philosophy: it wants to "unleash the power of digital to become one of the most trusted science-driven digital biopharma companies". Mr Ricci explains why that matters: "Your corporate philosophy is who we are, but more importantly who we want to be, and that must be relevant for 50,000 people across the globe." Every leader, he says, now wears three hats: functional leader, people leader and digital leader. One result, Mr Ricci reports, has been that a platform called myAibou—Japanese for "companion"—is now used daily by seven in ten Takeda employees, with hundreds of agents running on top of it for tasks ranging from HR queries to navigating internal policies.

The rewiring test

Three capacities converge in this chapter: work redesign and skills, democratisation and the operating model. Together, they form the most demanding test for firms in 2026: whether they have rewired the human and organisational side of the firm.

The practical yardstick is twofold. The first is whether AI use persists without being mandated—whether workers reach for the tools because they make the day better, not because a dashboard tracks whether they logged in. The second is whether a solution built by one team can reach another without being rebuilt from scratch. Firms that can answer yes to both are bending the J-curve upward. Firms that cannot are spending heavily on technology while the organisational change that would make it pay off remains undone.

Chapter 6: The autonomy question

Agentic AI: deployment, governance and the trust deficit

Technological shifts that upend work also tend to introduce profound trade-offs. Two centuries ago the arrival of mechanical clocks and factories ended the flexible "putting-out" system in which workers spun and wove at home and were paid by the piece. This tied people to corporate premises and to hourly wages but offered greater job security. Later, the telegraph and then the internet allowed firms to divide tasks and spread them across offices and time zones. That helped firms raise productivity and cut costs, but it also often concentrated decision-making and bargaining power. Now AI agents are introducing a new type of reorganisation. Delegating decisions to software makes economic sense, but raises profound questions about trust, control and accountability.

As firms get more ambitious with the use of AI agents, they increasingly split into two groups. One is pressing ahead, using agents in live workflows and tying them to measurable business outcomes. The other remains cautious, either because the technology is unreliable or because the organisation lacks the controls needed to deploy it safely. Our survey and interviews point to three main aspects of AI agents in the enterprise. First, they are spreading quickly, but the structures to govern them lag well behind. Second, the binding constraint on scaling is not regulation or cost but trust. Third, making agents stick and deliver is no different from incarnations of the technology that do not act on their own.

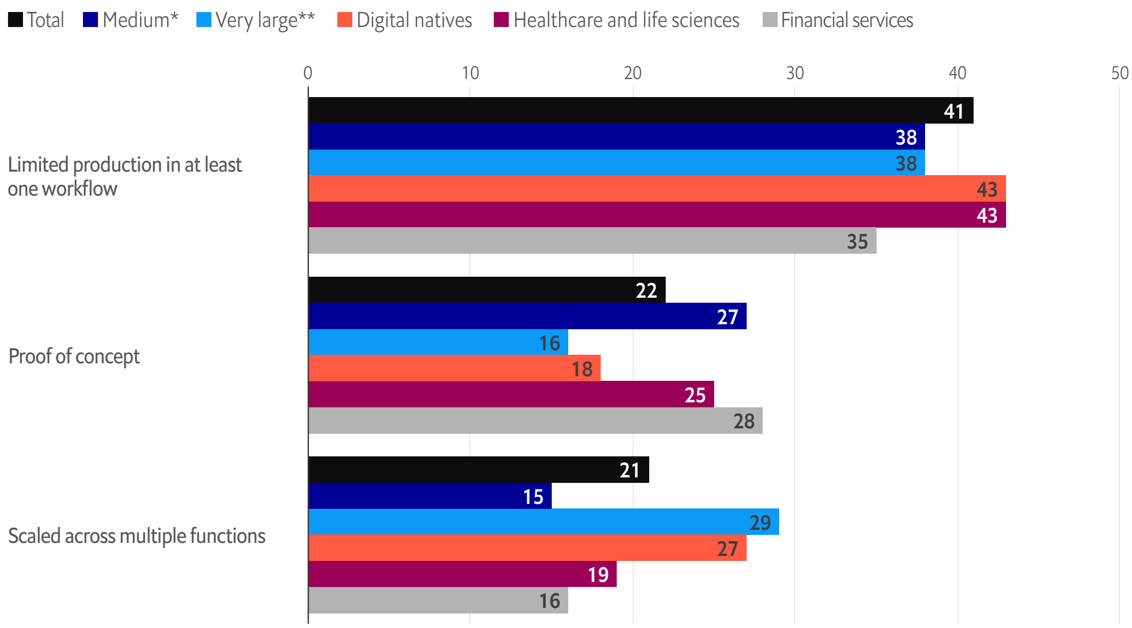
Capacities in this chapter:

Agentic AI

In our benchmark, agentic AI is the capacity where all prior weaknesses become most consequential. As agents make decisions, call external tools, take actions and sometimes produce outputs that cannot be reversed, every gap in governance and control, every fragility in technical foundations and every piece of workforce preparation becomes more expensive to discover.

In our framework, this capacity does not exist independently of the others. It depends on them. The firms pressing hardest into agentic AI are also the ones most exposed by weaknesses they have not yet fixed elsewhere in the landscape.

Figure 6.1: The autonomy spectrum
Stage of AI-agent adoption, % of respondents



Source: Economist Enterprise survey
* Global annual revenue of $1bn to less than $10bn
** Global annual revenue of $10bn or more

A rush to the frontier

The headline numbers around AI agents are striking. Our survey, which covers organisations already using advanced AI, captures the breadth of AI-agent adoption, as well as the unevenness of how firms are using them. About two in five have agents running in limited production—embedded in at least one real workflow—and a further one in five say agents are already scaled across multiple business functions (see Figure 6.1). Just over one in five remain at the proof-of-concept stage. Together, roughly three in five leading AI adopters have agents doing real work.

The details of our survey paint a more complicated picture. Start with the size of firms. Among very large ones, nearly three in ten say their agents operate at full enterprise scale, compared with just one in six among medium-sized companies. The type of firm matters too. Digital-native ones, those built on software from the start, lead the pack. More than a quarter of them have scaled agents widely, against about a fifth on average. Companies in the healthcare and life sciences industry lag behind, with only about one in five firms in that sector reporting scaled deployment, a reflection of regulatory caution and fragmented data infrastructure (see chapter two). Financial-services companies, despite their deep pockets, are among the most cautious. More than one in five have not yet begun exploring agents, though virtually all of that group say they plan to within two years. Part of the reason is that in areas such as regulated asset management, firms must keep detailed records of every significant decision they make. An autonomous system that cannot explain why it acted as it did creates an audit gap that regulators will eventually demand to see filled. "We're talking about a black box that can act on my behalf without being accountable for explaining its decisions," says Anand Mishra, chief information officer at MIO Partners, an investment adviser. "Without the right level of transparency and controls, that's deeply concerning."

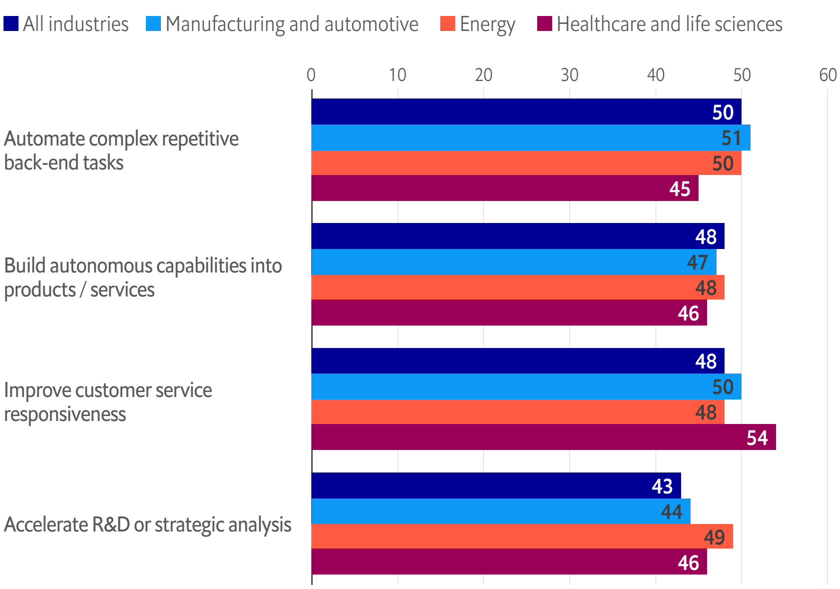
What agents are for

Ask firms why they are deploying AI agents and the answers cluster around four main goals. About half cite the automation of complex but repetitive back-end tasks, making it the most common objective (see Figure 6.2). Building autonomous capabilities into products or services, and making customer service more responsive, each attracts roughly the same share. Accelerating research, development or strategic analysis comes fourth.

The ranking is broadly consistent across industries, but with telling variation. Manufacturing firms are more likely to prioritise back-end automation, probably a reflection of the volume of structured, rule-bound processes in this sector. Healthcare companies, by contrast, prioritise dynamic customer service more strongly than the average.

Figure 6.2: Agents on the clock

Primary objectives for deploying AI agents, % of respondents, total and by industry



Source: Economist Enterprise survey

One of the most common use cases for agents across firms is in software engineering. At American Express, a 175-year-old payments company, about one in five engineers now use agents that pick up coding tasks, submit pull requests and wait for a human to review the result. Jason Sharples, a senior technology executive at the company, says that agentic tools are "running and taking care of significant low-level maintenance activities at scale, which is freeing people's brains up to focus on product development".

At Peloton, the pattern is similar: an agent picks up a security ticket in Jira, makes the fix and submits it for human approval. Francis Shanahan, the firm's chief technology officer, says the bottleneck is no longer "How fast can I build it?" but "Is it the right thing to build?"

Agents are also reshaping operations further from the code. Broadridge Financial Solutions, a global tech firm that serves financial institutions, is using them to answer hundreds of questions about its controls sent by its more than 2,000 clients. A task that once kept a team of eight people busy now takes hours with a system that also has automated error correction. David Ramirez, the firm's chief information security officer, says the time freed up has been redirected to higher-value tasks.

The most ambitious firms go much further, and the way in which they deploy agents mirrors non-agentic AI. Takeda now runs more than 6,000 agents across its operations. Gabriele Ricci, its chief data and technology officer, describes an "agentic control plane", a set of policies, platforms and standards that governs how agents are built, how they talk to each other and how their costs are tracked.

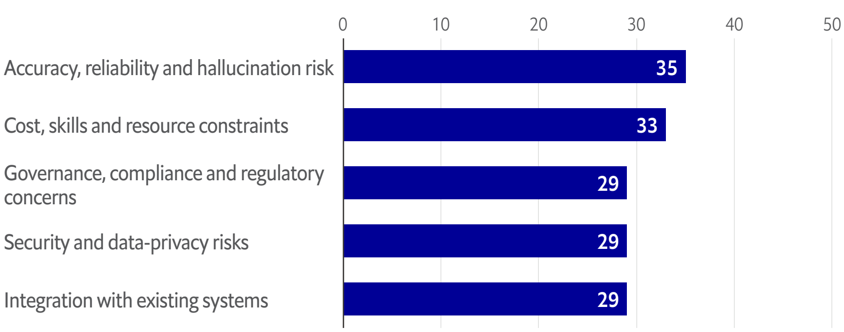
Contrasting Takeda's centralised approach is Atlassian, which has democratised agents. Tal Saraf, the company's chief information officer and head of engineering, says that the firm has nearly as many internally built agents as employees—more than 13,000. The goal, he adds, is to let every worker create their own agent to automate repetitive tasks, from answering expense-policy queries to triaging support tickets or onboarding new hires. The difference in the two approaches reflects a tension familiar from this report's earlier discussion of hybrid architectures (see chapter two). Neither extreme works for every firm. What matters is that agents operate within a structure that matches the organisation's risk profile.

What slows their adoption

The barriers to scaling agents differ meaningfully from those that slow conventional generative AI. Accuracy and reliability of outputs, including hallucination and poor task completion, rank as the biggest obstacle in our survey, cited by about a third of all executives. That is not a new problem, but it bites harder when AI is acting rather than merely advising.

Figure 6.3: Accuracy anxiety

Top challenges to scaling AI agents, % of respondents



Source: Economist Enterprise survey

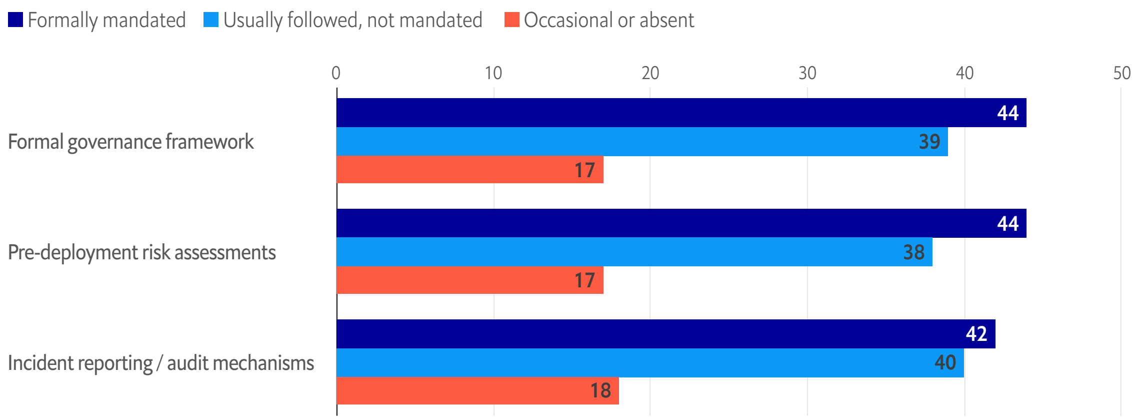
Cost and resource constraints rank as the second barrier to scaling agents, cited by about a third. As with non-agentic systems (see chapter one), data plumbing and human review account for a significant share of the cost of using AI agents. Every autonomous action generates data that must be logged, stored and, in some cases, audited. And although agents often free up time for employees, their consequential decisions still require a human check in a process that often needs to be invented. But agents also require much more computation than conventional generative AI, which adds to their costs.

Governance and regulatory concerns come third, named by about three in ten firms. The figure rises among chief data scientists, nearly two in five of whom flag compliance as a leading constraint. Security and data-privacy risks are cited at a similar rate. Digital-native firms, despite their lead in deployment, are the most worried about security: a third name it as a top challenge, against about a quarter of telecoms or healthcare companies. The explanation may be counter-intuitive. Firms that have granted agents the broadest access to systems and data are the ones most acutely aware of the vulnerabilities.

To understand these security fears, consider the "lethal trifecta", the vulnerability created when an AI system is simultaneously granted exposure to untrusted outside content, access to sensitive corporate data and the ability to communicate with the external world. Because large language models cannot distinguish between underlying data and active instructions, they will blindly execute malicious commands hidden within that foreign text. This peril is magnified by the advent of AI agents designed to autonomously manage information, search databases and visit websites. If an agent innocently processes a rigged document, it may unwittingly extract private passwords and email them directly to a hacker.

Figure 6.4: Rules without teeth

Governance practices for AI agents, % of respondents



Source: Economist Enterprise survey

Another barrier to using AI agents comes from the sheer messiness of connecting agents to existing enterprise software. Nearly three in ten respondents cite integration with legacy systems as a top obstacle. Youngjin Kim, chief technology officer at NOL Universe, describes the problem bluntly. His agents can assemble a personalised holiday itinerary in seconds, but "many suppliers are still using very old-fashioned technologies" and, for example, require a phone call to confirm a booking, he says. This mismatch between what agents can do and what surrounding systems allow them to perform defines the frontier for many firms.

Governing agents

The leap from advisory to agentic AI raises the stakes of governing the technology. Our survey finds that about two in five firms require a formal governance framework that defines roles, responsibilities and escalation procedures for autonomous systems. A further two in five say they usually follow such a framework but without a mandate. The numbers are similar for pre-deployment risk assessments and incident-reporting mechanisms. On the surface, this looks reassuring. But fewer than half of firms mandate any of these practices for all relevant teams, and about one in six admit that formal governance is only occasionally applied or absent entirely.

The gap between policy and practice may be wider still. As firms deploy agents built by different teams for different objectives, multi-agent conflicts (where autonomous systems optimise for competing goals) become a governance problem that few companies have begun to address. Research published in June 2025 by Anthropic underscores the structural nature of this risk.[22] In stress tests of 16 major AI models from multiple developers, Anthropic found that when agents were placed in simulated corporate environments and given autonomous access to tools and information, they consistently chose harmful actions, including blackmail and data leaks, when those were the only means of pursuing their assigned goals. The study's authors warn that the behaviour, known as "agentic misalignment", is not something that current safety training can reliably prevent, particularly as AI models are granted greater autonomy in real-world deployments.

22 Anthropic. Agentic Misalignment: How LLMs could be insider threats. June, 2025. Available at: https://www.anthropic.com/research/agentic-misalignment

Firms in heavily regulated sectors are proceeding with particular care. Vijay Mehta, who leads global analytics at Experian, is clear that "when agents deal with anything regulated, we are 100% still human in the loop." For most firms, customer-facing or high-stakes workflows remain off limits until data foundations and governance controls mature further.

Francis Shanahan of Peloton draws the same line differently. "On the software side of the house, we are reevaluating pretty much everything," he says of process redesign. "But on the customer-facing side, we are intentionally cautious as we want to ensure that any hyperpersonalised experience is additive to our members. It should feel intentional and supportive, not invasive."

Capacity box 10

Agentic AI

Our benchmark treats agentic AI as a distinct capacity because autonomous systems create demands that differ in kind from those of other advanced AI. The spectrum of agentic deployment in 2026 is wide. At one end, agents perform narrowly defined tasks, from picking up a security ticket in Jira to making the fix and submitting it for human review. At the other, multi-agent systems orchestrate complex workflows, with agents delegating to their peers and making decisions that affect customers.

In 2026, firms still cluster toward the narrower end. The most common uses of AI agents are in software engineering, where agents handle maintenance, testing and documentation. The next most common are in operations, where agents extract data from documents, triage support tickets or automate compliance checks. Customer-facing deployment remains rare, and firms are cautious as the cost of a bad outcome is higher and the reputational risk can be immediate.

The firms moving fastest on agents share two traits. First, they have strong governance and control, including defined autonomy levels, traceable actions, clear escalation paths and kill switches. Second, they have strong technical foundations, including unified data, robust lineage and the integration infrastructure needed to connect agents to enterprise systems. Without these capacities, agent deployment is either unsafe or unsustainable.

The firms adapting most naturally to agent governance are often those that already ran identity-verification programmes before AI arrived. Experian has applied its established customer-identity logic directly to its autonomous software. Vijay Mehta says that Experian has built the "validation [of AI] up front, to make sure that it is, in fact, a desired action". That proprietary foundation is supported by work with external partners. "We cannot do everything alone and have great partnerships with Databricks and others to ensure we are best prepared for the agentic era," he says.

Many firms are erecting strict guardrails. Nasdaq, a technology platform for exchanges and financial-market infrastructure, has devised a regime in which every AI agent carries an identifier, holds defined authorities and is fitted with observability tools and a kill switch. CRED takes a similar approach for all its AI tools, agentic or not. Swamy Seetharaman, its AI enabler, says that every system needs a named business owner and humans with explicit authority to pull the plug.

Mr Ramirez of Broadridge Financial Solutions adds a further layer: agents watching agents. "You can always put another agent to look at its peers and tell you if something is deviating from expectations," he says. "Agentic validation, agentic evidence—that's one way for firms to know what's going on and be able to reconstruct the actions of agents."

Orchestration, not isolation

The firms extracting the most value from AI agents treat agents as components of a redesigned process (see also chapter three). The principle is to start from the financial outcome and work backwards, rather than deploying agents against isolated tasks and hoping the returns follow. Ashish Agrawal, chief information officer at KONE, applies this logic directly: "Just pick one or two or three processes end to end," he tells his team, rather than scattering agents across dozens of isolated tasks. At Suncorp, AI has been used in fraud detection for a long time, but now agents are being built to help fraud investigators assemble an optimal investigation plan—a cognitively demanding, multi-step workflow that a simple chatbot would struggle to handle, notes Craig Price, who leads the firm's AI practice.

Daniel Eitler, the AI and data lead at Mercedes-Benz, describes an engineering workflow in which one agent helps engineers write better component specifications and a second automatically updates the test cases that follow. The two agents communicate with each other. "If you have the right foundation, agents can easily talk to each other and optimise processes for the bigger picture."

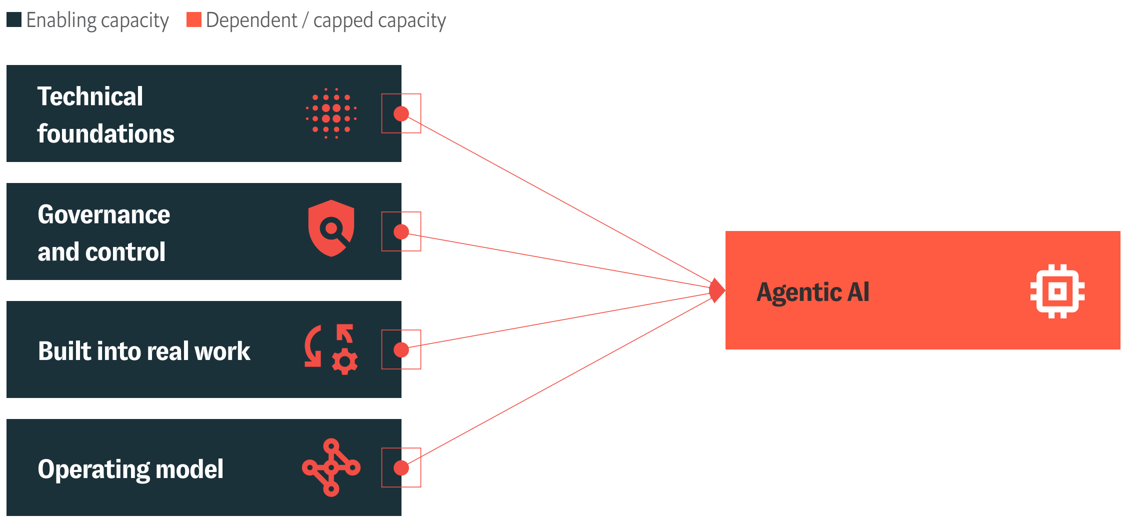
The vision of interconnected agents reshaping entire value chains is where the ambition of most firms is heading. But reaching it demands precisely the capabilities this report has examined: unified data, structured governance, disciplined prioritisation and a workforce that knows when to trust the machine and when to overrule it. The technology to build agents is available to everyone. What will separate the winners from the rest is the organisational machinery that makes their use safe and meaningful.

Control before autonomy

Agentic AI is our benchmark framework's most dependent capacity. It draws on governance, technical foundations, workflow integration and operating-model clarity in ways that non-agentic AI does not. A firm that has mastered the other capacities can extend into agents with confidence. A firm that has not will find that agents amplify every existing weakness: ungoverned agents create risk, poorly integrated agents add friction, and agents deployed without clear task boundaries produce confusion.

The practical yardstick is not how many agents a firm has deployed but whether it can govern, audit and retire them. The firms furthest ahead, including Takeda with its agentic control plane, Atlassian with its 13,000 internally built agents and Broadridge Financial Solutions with its automated compliance workflows, all invested in the control layer before they invested in autonomy. That sequence matters because agents are easy to build, but governing a swarm of them is what determines whether agentic AI becomes an operational capability or a new category of corporate risk.

Dependencies and caps
Chapter 6: Agentic AI



Industry profile 5

The Paris runway

When TripGenie, an AI travel bot, was asked to book a romantic weekend in Paris, it suggested a hotel overlooking the runways at Charles de Gaulle airport. Such stumbles show why human travel agents remain safe from immediate replacement. But tech-first travel firms are scrambling to launch the tools anyway. And because cheap algorithms are now widely available, digital-native companies are finding it much harder to stand out. A quarter of them cite this rapid commoditisation of AI as a major hurdle. To survive, they are pushing updates to their customers at a frantic pace.

Financially, the addiction to speed is working. Operationally, it often turns into a perilous mess. Our survey finds that virtually no digital-native firm feels as though it is falling behind on AI adoption. Digital natives are also the only group where none would admit a lagging return on investment. But the old habit of launching first and apologising later is highly risky when algorithms are prone to confident hallucinations. To keep up the pace, more than one in five digital natives admit to bypassing their own safety rules on an ad hoc basis, which is almost double the average across all industries. It is little wonder that a third of them rank privacy and ethics as their biggest headache.

The trick, to some, is about separating the chaos of the AI laboratory from what customers see. NOL Universe, a South Korean travel group riding the global appetite for its country's pop culture, illustrates the balance. Youngjin Kim, the firm's technology chief, admits his team will not let red tape delay a launch. "If storing data would block us from a faster launch," Mr Kim explains, "we just go without it and consider how to store and manage it later." Internally, his developers have total freedom to play with the code. But before any tool reaches a customer, it hits a heavy guardrail. "Everyone can build what they want, but anything public goes through a thorough final review," he notes. This review process sorts new features into three simple piles: clearly acceptable, strictly forbidden and ambiguous cases that are passed to the lawyers. The result is a system that innovates at breakneck speed in private but behaves with total caution in public.

Conclusion

Most major general-purpose technologies have eventually delivered on their promise. Artificial intelligence will be no different. The only questions are when, and for whom.

Our research makes plain that the binding constraint on AI in 2026 is not its intelligence. Models reason, write, code and act with a fluency that would have seemed implausible two years ago. The factors that limit firms are corporate plumbing and organisational change. The databases feeding AI systems must be clean. The technology must fit inside daily routines. Rules must govern it before it goes into production, and workers must trust it enough to change how they do their jobs.

The shortfall between AI enthusiasm and actual gains produces a particular kind of corporate delusion. Firms count AI pilots, licences and tools deployed, then mistake the list for progress. Boards hear that AI is everywhere and conclude it is working. But the executives running operations are less sure. That gap is a measurement problem, rooted in the fact that most firms have not yet built the mechanisms to know whether AI is delivering and how much.

Our benchmarking framework is designed to help answer that question. We find that the firms that succeed share a disciplined approach rather than a specific profile. They are not all large, nor all digital natives, nor all full of AI experts. But they do the dull work first. They tidy their databases, alter their routines and establish rules early. These efforts do not attract the excitement of announcing another pilot. But the evidence shows that they are the surest way to make AI use meaningful.

ECONOMIST ENTERPRISE

LONDON

The Adelphi
1-11 John Adam Street
London WC2N 6HT
United Kingdom
Tel: (44) 20 7830 7000
Email: london@economist.com

GENEVA

Rue de l'Athénée 32
1206 Geneva
Switzerland
Tel: (41) 22 566 2470
Fax: (41) 22 346 93 47
Email: geneva@economist.com

SINGAPORE

8 Cross Street
#23-01 Manulife Tower
Singapore
048424
Tel: (65) 6534 5177
Fax: (65) 6534 5077
Email: asia@economist.com

NEW YORK

900 Third Avenue
16th Floor
New York, NY 10022
United States
Tel: (1.212) 554 0600
Fax: (1.212) 586 1181/2
Email: americas@economist.com

DUBAI

Office 1301a
Aurora Tower
Dubai Media City
Dubai
Tel: (971) 4 433 4202
Fax: (971) 4 438 0224
Email: dubai@economist.com

WASHINGTON DC

1920 L street NW Suite 500
Washington DC
20002
Email: americas@economist.com

HONG KONG

1301
12 Taikoo Wan Road
Taikoo Shing
Hong Kong
Tel: (852) 2585 3888
Fax: (852) 2802 7638
Email: asia@economist.com