Institute for DIGITAL INNOVATION & AI

THE UNIVERSITY OF MELBOURNE

Melbourne Business School

AI and the Redesign of Work

How CxOs, boards, policymakers, and educational institutions should govern AI's structural impact on work, talent, and organisational performance.

Drawing on a roundtable of Australian CEOs, Chief People Officers, CTOs, Non-Executive Board members, university professors, and the current research evidence

Table of Contents

Executive Summary

The public debate about AI and employment keeps rehearsing the same arguments.
One side says AI is hollowing out knowledge work and wrecking the graduate job market. The other says displacement is overstated, employment remains resilient, and history shows that transformative technologies ultimately create more jobs than they destroy.

Both are answering the wrong question. The debate offers executives and policymakers a choice between alarm and reassurance. Neither tells them what to actually do, and neither matches the pace at which the shift is already happening.

The right questions are structural: what AI is doing to how tasks are composed, how capability is built, how value is created and sustained, and then what that means for how organisations need to be redesigned and governed. The deeper question is whether organisations are making those choices deliberately or by default. The costs of ungoverned transition will not appear on the balance sheet for five to ten years, by which point they will be very difficult to reverse. Rather than viewing it as just a governance failure at the organisational level, we should view it as a choice at the level of society, possibly made by default, about the kind of work, capability, and economic life we want on the other side of this transition.

> ***AI poses a commercial, financial, and societal governance challenge. The labour market disruption is real, but it is a consequence not the main game. The main game is all about which organisations build durable human capability and how, and which quietly hollow it out, and the choices we make as a society about how we engage with this technology. If we treat it only as a jobs story, we risk missing what really matters.***

Three propositions should sit at the centre of every AI strategy discussion. First, AI is removing the work through which organisations have historically built their leaders, creating what we call human-capital debt: work is produced, but the worker cannot explain, challenge, or improve upon it. Second, without governance, AI *industrialises the appearance* of productivity rather than its substance. The output is produced, but the underlying thinking may not have occurred. Third, technical capability is not the same as economic disruption. The gap between them is closing at the rate organisations choose to govern it.

This paper draws on a roundtable of Australian business leaders, policymakers, and business school leaders set against the current research evidence. It is addressed to four audiences with the most leverage over outcomes: chief executives and their leadership teams; boards and non-executive directors; national and regional policymakers; and universities and business schools. For each audience, the challenge is different.

Organisations that treat AI as a cost-reduction tool risk weakening their capability base and leadership pipeline. Those that treat it as an operating model redesign opportunity by reshaping roles, decision rights, and developmental pathways are more likely to build durable advantage. The critical implication is that if organisations do not redesign work deliberately, that redesign is likely to occur by default.

SECTION 1

The Problem with the Current Debate

In May 2025, Dario Amodei, co-founder of Anthropic and one of the people with the clearest view of where frontier AI is heading, told Axios that "artificial intelligence could eliminate half of all entry-level white-collar jobs and push unemployment to between 10 and 20% within one to five years." He accused his own industry and the political class of "sugar-coating" what was coming, and named the sectors most at risk: technology, finance, consulting, and law (VandeHei and Allen, 2025).

Months later, Klarna's CEO Sebastian Siemiatkowski told Harry Stebbings on the 20VC podcast that his company's workforce would shrink from 3,000 employees today to under 2,000 by 2030, cutting roughly a third, not through layoffs but through natural attrition of around 20% a year combined with a freeze on replacement hiring. Siemiatkowski explicitly aligned himself with Amodei, saying "I'm more in Dario's camp. I want to be honest about the fact that I do think there's going to be a very big shift" (Munis, 2026). By early 2026, Jamie Dimon, CEO of JPMorgan Chase, was telling investors at a company event that "...now's the time to start thinking about what you do..." if AI-driven job displacement becomes a problem. He added that JPMorgan "wasn't going to put its head in the sand" and that the bank "would likely employ fewer people in the next five years" thanks to AI-driven productivity gains (Quiroz-Gutierrez, 2026). The warning is not confined to US technology and Wall Street voices. In May 2026, Commonwealth Bank chief executive Matt Comyn used a contributed piece for the *Australian Financial Review*, timed to CBA's first Accelerate AI conference, to argue against soft-pedalling the workforce disruption. For Comyn, false reassurance does not protect workers; it only ensures they are surprised later.

The counter-chorus has been just as loud. Morgan Stanley's chief U.S. economist Michael Gapen recently argued that fears of AI-driven job losses "appear overstated," with evidence so far pointing to limited disruption, consistent with historical patterns, followed by rising productivity, much as in every previous general-purpose technology wave (Morgan Stanley, 2026). Goldman Sachs Research estimated that generative AI would raise labour productivity by around 15% when fully adopted, with unemployment rising by roughly half a percentage point during the transition and the effect largely fading within two years (Goldman Sachs, 2025). The Budget Lab at Yale, analysing post-ChatGPT labour market data through early 2026, finds no clear relationship so far, in aggregate data, between AI exposure, automation, or augmentation measures and changes in employment or unemployment. While the occupational mix is evolving, those shifts largely predate generative AI and remain within historical ranges, suggesting that any labour market effects have yet to register in aggregate data (The Budget Lab at Yale, 2026). Anthropic's own research team, proposing a new "observed exposure" measure built from real Claude usage and applying it to US labour market data, found "no systematic increase in unemployment for highly exposed workers" since late 2022 (Massenkoff and McCrory, 2026), a striking early finding from the same company whose CEO had warned of a bloodbath.

And then there is the scenario nobody wants to model. In February 2026, the financial research firm Citrini published *The 2028 Global Intelligence Crisis*, a fictional macro memo written from the perspective of June 2028 describing how an initial wave of white-collar layoffs in 2026 compounded into a feedback loop. AI improves, companies lay off workers, savings fund more AI, displaced workers spend less, consumer-facing firms weaken, margins compress, firms invest more in AI - all with "no natural brake." The piece imagined unemployment at 10.2%, the S&P 500 down 38% from its 2026 highs, and labour's share of GDP collapsing from 56% to 46% in four years (Citrini and Shah, 2026). It was explicitly labelled a scenario, not a prediction.

It went viral anyway, because it articulated the left-tail risk that the reassurance camp has no clean rebuttal for. Executives and boards are now left choosing between these narratives, and the choice tends to track whichever confirms the prior beliefs they already held. Neither position tells a CEO what to do on Monday morning.

> ***When the CEO of the company building the technology, the CEO of the largest US bank, and the CEO of one of Europe's largest fintechs are all publicly warning that something is coming, and the economic data show almost nothing has happened yet, the only interesting question is why the gap exists, and what closes it.***

Our view is that neither side is really asking the right question. The alarm camp translates technical capability into negative labour-market outcomes. The reassurance camp leans on historical averages from transitions that unfolded over decades, not quarters, and perhaps does not recognise that this technology might be fundamentally different from what we have seen historically. Neither tells us how to govern an organisation through the transition, and neither takes seriously the possibility that the technology is simultaneously being over-hyped in its near-term displacement effects and underappreciated in its structural ones that will affect the long-term. Roy Amara's well-known technology law - "We tend to overestimate the effect of a technology in the short run and underestimate its effect in the long run" - rings true in today's AI age.

The foundational error on the alarm side is the conflation of exposure with displacement. These are distinct analytical categories that the public debate collapses into one. Exposure measures which tasks a technology could in principle affect. In contrast, displacement measures which workers actually lose their jobs. Eloundou, Manning, Mishkin, and Rock estimate that "around 80% of the US workforce could have at least 10% of their tasks affected by AI," with approximately 19% facing exposure to more than half of their tasks once complementary software tools are included (Eloundou et al., 2024). Around 15% of tasks could already be completed significantly faster through direct LLM use, rising to between 47 and 56% when purpose-built applications are considered. But these are exposure estimates, not displacement forecasts. The gap between the two is driven by social norms, regulatory frameworks, organisational inertia, the cost of complementary systems, and managerial/investor/shareholder choices.

The reassurance camp makes the mirror error. It points to the aggregate labour market (e.g., unemployment near historical norms, payrolls growing, hiring continuing) and concludes that the structural story is overblown. But aggregate employment figures are structurally incapable of detecting a hollowing-out of entry-level work until it has already happened. The hiring rate in the US fell to 3.1% in February 2026, the lowest since April 2020, even as headline unemployment stayed benign (US Bureau of Labor Statistics, 2026). A hiring freeze shows up in layoff data years after the career ladder it breaks is already gone. The reassurance position is not wrong about today, but it may be wrong about where today is pointing.

There is a further problem the headlines miss. The economic outcomes of AI diffusion are not determined by the technology alone. They depend on who captures the rents across the AI stack, how the technology is priced into markets, and how organisations and regulators respond. Productivity gains can coexist with falling real wages for some workers if rents are captured upstream. That is a governance question, not a forecasting one.

A second, less visible error is measurement. Much of AI's economic impact is intangible and expensed rather than capitalised, which means early evidence looks weaker than the underlying transformation is. We return to this in Section 2. For now, it is enough to note that both the alarm and reassurance camps are reading numbers that systematically understate what is happening inside firms.

A more useful framework separates four stages: exposure (which tasks AI can affect in principle, per Felten, Raj, and Seamans, 2019; Eloundou et al., 2024), adoption (which workers adopt and organisations deploy it; see Bick, Blandin, and Deming, 2025; Hartley et al., 2025), substitution (which tasks are genuinely replaced), and complementarity (which tasks are augmented or newly created).

This four-stage framing aligns with recent work cautioning that the public debate collapses several distinct measures into one. Floridi, Novelli, and Morley (2026), auditing twenty-six prominent forecasts, identify six quantities that are routinely treated as interchangeable - exposure, automation potential, cost-effective automation, displacement, net employment change, and productivity uplift - and argue that naming which one is being reported is a precondition for any claim that can be tested at all.

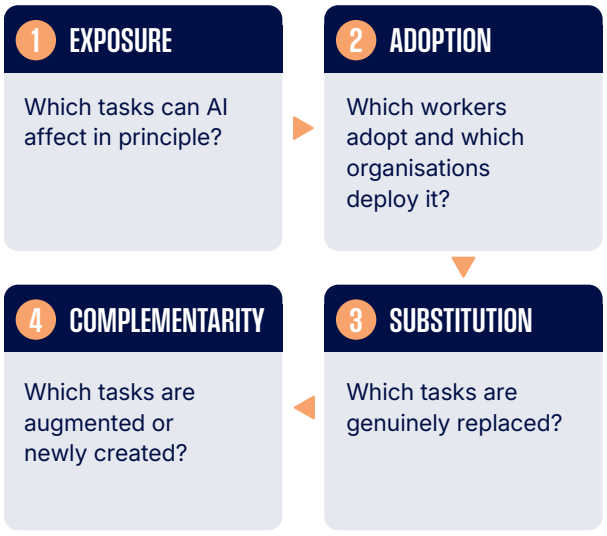


Figure 1. Four-Stage Framework of Exposure, Adoption, Substitution, and Complementarity

Mäkelä and Stephany (2024) provide the clearest empirical evidence to date that the latter two dynamics coexist: in 12 million US job postings, AI-intensive roles are 22% more likely to demand complementary skills and 34.8% less likely to demand substitutable ones, and in the most exposed industries, the spillover produces around 1.7 complementary roles for every substitutable role lost.

Exposure is not displacement, and adoption is not substitution. Organisations that confuse these categories will mismanage the transition, either by underinvesting in AI capability or by over-indexing on headcount reduction at the expense of institutional resilience.

This is why the Amodei-versus-Yale divide is a false one. Amodei is largely right about capability. The Budget Lab is largely right about aggregates *so far*. Both can be true because the mechanism running between them - adoption, substitution, redesign of work - is where the action is, and that mechanism is governed by choices organisations and policymakers are making right now. The useful question for a CEO, a board, or a regulator is simply: "what is the structure of work we are building, by design or by default, as this technology diffuses?"

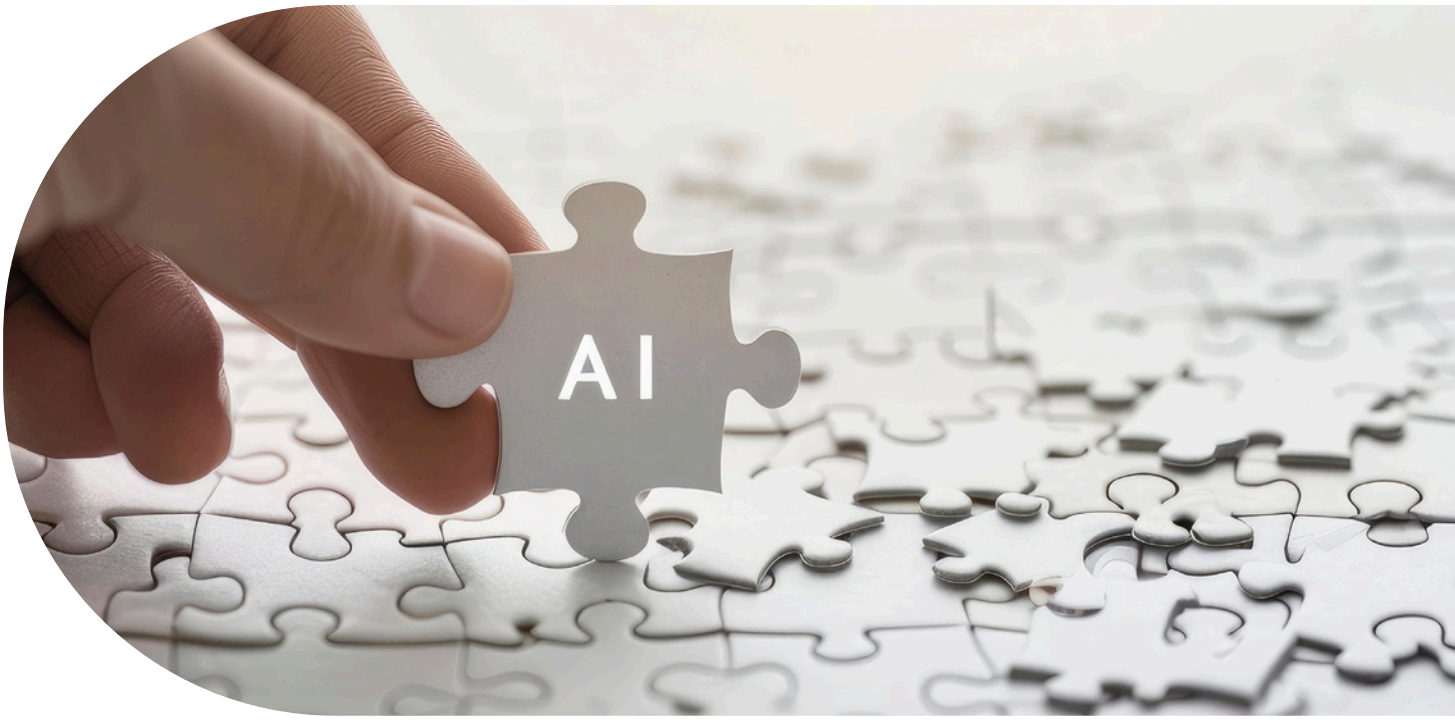
Why capability does not equal diffusion

The most common error in the alarm narrative is the assumption that because AI capability is compounding, economic disruption will compound at the same rate. That assumption is precisely what drives scenarios like the Citrini memo, and precisely what Citadel Securities took aim at in its response. Citadel's macro strategist Frank Flight argued that the current debate "conflates the recursive potential of the technology with expectations of recursive economic deployment" (Citadel Securities, 2026). Technical capability can scale in quarters. Organisational adoption historically follows an S-curve: slow and expensive early, accelerating as complementary infrastructure develops, then plateauing as saturation and integration costs bite. Markets routinely extrapolate the acceleration phase linearly. History says clearly that the plateau arrives.

Flight's second argument is physical. If automation expands rapidly, compute demand rises, which pushes up the marginal cost of compute. Once that marginal cost exceeds the marginal cost of human labour for a given task, substitution stops, creating what Flight calls "a natural economic boundary" that narratives assuming frictionless replication of intelligence simply ignore (Citadel Securities, 2026). Exposure reflects what AI could do. Diffusion reflects what it does do, bounded by capital, energy, regulation, and organisational change. Collapsing the two is how serious people end up forecasting 20% unemployment on a five-year horizon.

None of this means the structural shift is not real. The executives closest to the technology are not wrong to be worried.

When Amodei, Dimon, and Siemiatkowski are publicly saying the same thing, something is happening. It means the time horizon is longer, the mechanism is more contingent on organisational choice, and the appropriate response is governance. *The Wall Street Journal* recently asked five leading AI executives, including Anthropic's Daniela Amodei, what they advise their own children to study. Their answers consistently emphasised adaptability, judgement, communication, and distinctly human capabilities, and the capacity to direct AI rather than compete with it (Weber, 2026). Section 2 turns to what the evidence actually says about where human value is migrating, and whether the advice these executives are giving their children matches what the research shows.



SECTION 2

What the Evidence Actually Shows

The research literature on AI and work has matured considerably in the past three years, spanning experimental studies, firm-level analyses, and international policy assessments (Gmyrek, Berg, and Bescond, 2023). A more granular picture is emerging, one that challenges simple narratives of both displacement and immunity, and points instead toward a jagged and differentiated pattern of impact. A critical implication of this emerging evidence is that large language models exhibit the characteristics of general-purpose technologies: they are pervasive across tasks and improving over time, and they unlock a wide range of complementary innovations. Their economic impact therefore depends less on the models themselves and more on the co-invention of systems, processes, and organisational redesign that enable their effective use.

This also helps explain why measured productivity gains often lag technological capability. Historical evidence on general-purpose technologies shows that performance improvements depend on complementary investments in organisational capital. In the case of AI, these include data infrastructure, evaluation systems, and workflow redesign. Because these investments are largely intangible and expensed rather than capitalised, their contribution is difficult to capture in standard metrics. This creates the impression that AI's impact is slower or weaker than it actually is, particularly in the early stages of adoption.

This pattern is consistent with earlier evidence on information technology as a general-purpose technology. Bresnahan, Brynjolfsson, and Hitt (2002) show that productivity gains from IT adoption depended critically on complementary investments in workplace organisational changes, including decentralised decision-making, new work practices, and skill upgrading. Firms that adopted the technology without these organisational changes realised significantly lower returns.

The implication for AI is direct: the technology itself is unlikely to generate sustained performance gains without corresponding redesign of work and organisational structure.

Augmentation and substitution coexist

Experimental evidence shows that AI improves performance on a range of knowledge tasks. Dell'Acqua and colleagues show that management consultants using AI achieved significantly higher scores on complex creative assignments, but also that those who over-relied on the tool in domains where it was unreliable performed worse than those who applied more critical judgement (Dell'Acqua et al., 2026). Noy and Zhang document productivity gains in professional writing tasks, with lower ability workers benefiting most (Noy and Zhang, 2023). Girotra and colleagues find that AI substantially expands the quantity and improves the average quality of ideas generated in product development contexts (Girotra et al., 2023).

The emerging field evidence reinforces this pattern but adds an important nuance. In one of the first large-scale studies of AI deployment in a real workplace, Brynjolfsson, Li, and Raymond (2023) analyse over 5,000 customer support agents and find that access to an AI assistant increased productivity by 14% on average, with gains of more than 30% for less experienced workers. The mechanism is instructive: the system appears to diffuse the tacit knowledge of high performers, enabling lower-skill workers to move more quickly down the experience curve. This suggests that AI can act not only as a productivity tool, but as a capability transmission mechanism, compressing differences in performance across workers.

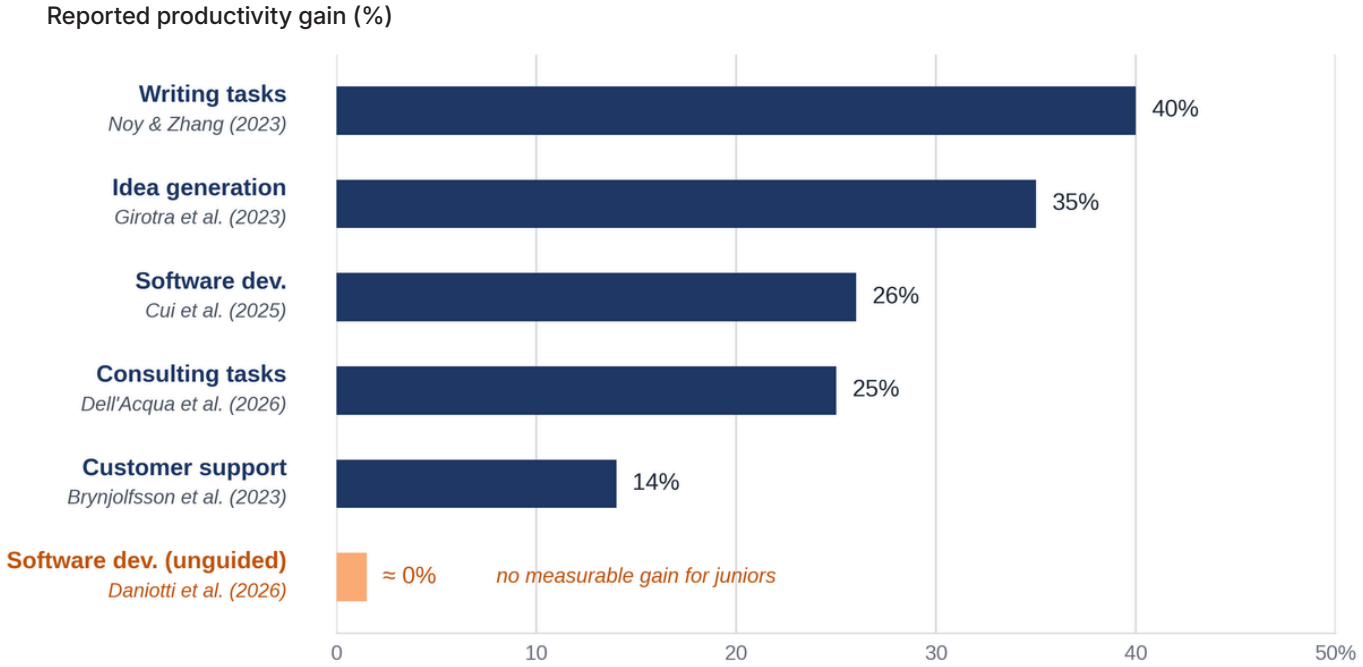


Figure 2. Estimated Productivity Gains by Profession across Multiple Studies

These findings matter, but they need careful interpretation. Performance gains in experimental settings do not automatically translate into headcount reductions. They may instead translate into higher output from the same workforce, higher quality from lower-cost workers, or expanded scope of what can be attempted. All of these create value without necessarily reducing employment. A further complication in interpreting these findings is endogeneity. Organisations that adopt AI in high-value areas are not randomly selected; they are often better managed and more digitally mature. Observed productivity gains may therefore reflect underlying organisational capability as much as the technology itself. This makes it difficult to infer causal impact from observational data and reinforces the importance of how AI is governed and integrated into workflows. Taken together, this body of evidence establishes that AI can improve performance at the task level. But it leaves open a critical question: how these gains translate beyond controlled settings into firms and markets.

Experimental designs hold the organisation fixed by construction. The more consequential question is what happens when firms reorganise work, expand output, and adjust their workforce in response to AI.

This point is reinforced by cross-country evidence showing that differences in productivity outcomes from similar technologies are largely explained by differences in management quality and organisational practices. Bloom, Sadun, and Van Reenen (2012) find that firms with more structured management practices and better organisational design extract significantly greater value from IT investments. Applied to AI, this finding suggests that observed performance differences across firms are likely to reflect variation in managerial capability and organisational readiness, not just access to the technology.

James Bessen's historical research at Boston University documents that industries experiencing rapid technological change often show employment growth rather than contraction when cost reductions stimulate sufficient demand. Automation in the textile industry drove down the cost of production in the nineteenth century, leading to a roughly 100-fold increase in cotton cloth consumption, and employment in the industry soared until the 1960s (Bessen, 2018).

Recent large-scale evidence provides that missing link. Using administrative data covering nearly all U.S. employers, Johnston and Makridis (2026) show that sectors with higher AI exposure experienced output increases of approximately 7-10% following the introduction of enterprise AI tools. Employment and wage bills also increased, but with an important distinction. Where AI augments human work, employment rises alongside output. Where AI can perform tasks independently, employment effects are small or insignificant *for now*.

This distinction is critical. It shows that productivity gains from AI do not mechanically translate into labour displacement. Instead, they expand output and, under many conditions, increase demand for labour (Makridis, 2026). The labour market impact is therefore not determined by exposure alone, but by how organisations structure work around augmentation versus substitution.

This pattern becomes clearer when we return to specific domains such as software development. Tracking nearly five thousand developers at Microsoft, Accenture, and a Fortune 100 manufacturer, Cui and colleagues found that AI coding assistance increased completed tasks by 26% on average, with gains concentrated among less experienced workers (Cui et al., 2025). A contrasting picture emerges from observational data: analysing more than thirty million public code contributions from 160,000 developers across six countries, Daniotti, Wachs, Feng, and Neffke found that early-career developers were the most enthusiastic adopters of AI coding tools, yet derived no statistically measurable productivity or capability benefit, while senior developers gained substantially in both output and technical breadth (Daniotti et al., 2026).

These findings about the productivity of junior versus senior developers are not necessarily in conflict. The productivity gains in the randomised trials arose in structured environments with defined tasks, training, and incentives, whereas the observational data reflect heterogeneous, unguided open-source adoption across a wide range of contexts.

Differences in measurement, task complexity, and user behaviour are therefore likely to explain part of the divergence. Taken together, the evidence suggests that productivity gains from AI are contingent on how it is deployed and integrated, rather than arising automatically from access alone (Makridis and Brynjolfsson, 2026).

Occupations are bundles, not monoliths

A critical error in popular analysis is the treatment of occupations as undifferentiated wholes. Occupations are bundles of tasks and skills, and AI's impact varies substantially across the components of any given role. Acemoglu and Autor's task-based framework shows that routine cognitive tasks are most susceptible to automation, while non-routine cognitive tasks, particularly those involving judgement, communication, and adaptability, are more durable (Acemoglu and Autor, 2011). Felten, Raj, and Seamans demonstrate that occupations requiring higher-level language and social skills are paradoxically more exposed to AI but also more capable of benefiting from AI augmentation (Felten et al., 2023).

The practical implication is that organisations need to think at the level of tasks, not job titles. A software engineer's role includes code generation, code review, system design, stakeholder communication, risk assessment, and team coordination. AI may accelerate or partially substitute for the first of these but it also strengthens the relative importance of the last four. The US Bureau of Labor Statistics reports approximately 1.7 million software developer jobs in the United States in 2024, with employment having continued to grow modestly since 2019, even as AI coding tools have proliferated. BLS projects a further 16% increase in this category by 2034 (US BLS Occupational Outlook Handbook, 2025; Meyersohn, 2026). Managing this shift requires role redesign. These task-specific effects have not, as yet, manifested in aggregate labour market data.

The same logic applies to other professions. One good example is radiology. Nobel Laureate Geoffrey Hinton famously said in 2016 that we should stop training radiologists because "within five years, deep learning is going to do better than radiologists... it might be ten years." And yet, we still have radiologists around in 2026, and in fact, we face a significant shortage of radiologists in Australia. Using the same logic as we used for software developers, radiologists read images, form judgements where there is ambiguity, talk with patients about those images, discuss images and diagnosis with other medical specialists, and train radiologists. The task that AI is best at is image recognition, which is one important component of radiologists' work. However, the role also includes diagnosis under uncertainty, communication with patients and clinicians, coordination with other specialists, and teaching. AI therefore changes the composition of work rather than fully eliminating the profession, increasing the relative importance of these higher-value tasks. This is a good example of AI freeing up time for a profession that allows them to put more resources into tasks for which humans are best designed. In other words, AI redefines a radiologist's work, perhaps making them more, not less, valuable. What is unclear at this stage about the future is whether (i) there will be a decline in the number of radiologists required per patient or (ii) there will not be a decline in the number of radiologists per patient but the value provided by each radiologist will, on average, improve.

The jagged frontier

Dell'Acqua et al.'s (2026) description of a "jagged technological frontier" is among the most useful contributions to the practitioner literature, and it deserves to sit at the centre of how leaders think about deployment. AI performs above human average on some dimensions, such as speed of drafting, pattern recognition, and consistency within defined rule sets. It performs below human average on others, including calibration of uncertainty, ethical reasoning, political judgement, and tasks requiring tacit social knowledge.

The frontier is not a line on a map that divides roles into above or below the line. It runs through every role and workflow, and through individual decisions. Workers who understand where the frontier runs outperform those who either distrust AI entirely or defer to it uncritically.

> *The organisations most at risk from AI are not the ones adopting it too slowly. They are the ones adopting it without the governance and design infrastructure to understand where it helps and where it misleads.*

Across experimental studies, firm-level evidence, and sector-wide data, a consistent pattern emerges: AI raises productivity, but its impact on labour is contingent on organisational design, task composition, and the balance between augmentation and substitution.

How organisations are actually responding

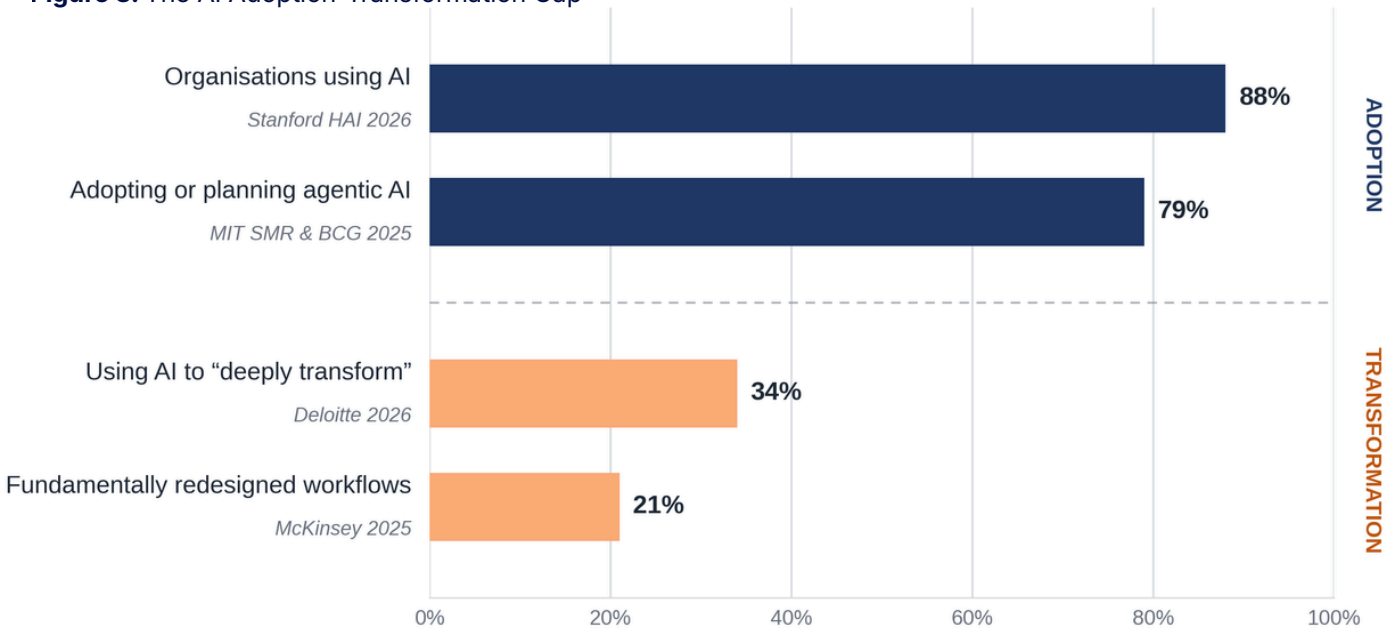
The contingent nature of AI's impact raises an empirical question: are organisations actually doing the work of deployment and organisational redesign that the evidence shows is required? Multiple independent enterprise surveys spanning academic, hybrid, and vendor sources converge on the same answer: adoption is near-universal, but transformation is rare.

Stanford's 2026 AI Index Report records 88% organisational AI adoption, with one-third of employers anticipating workforce reductions in the coming year and employment for software developers aged 22 to 25 down nearly 20% from 2024 (Stanford Institute for HAI, 2026). MIT Sloan Management Review and BCG find that 79% of organisations are adopting or planning to deploy agentic AI, but few have restructured workflows, governance, or talent strategies to match. Their summary line: AI adoption is racing ahead of strategy (Ransbotham et al., 2025).

Deloitte's 2026 enterprise survey of 3,235 executives finds only 34% are using AI to "deeply transform," while 37% are using it "at a more surface level, with little or no change to existing processes"; just 33% are redesigning career paths in response to AI, and only 30% are reimagining organisational structures (Deloitte, 2026). McKinsey identifies just 21% of organisations as having fundamentally redesigned workflows (McKinsey, 2025). PwC reports that even among organisations adopting AI agents, less than half are rethinking operating models or redesigning processes from ground up (PwC, 2025).

The empirical pattern that seems to be appearing is that AI deployment is widespread but organisational redesign in response to it is not. The gap between what the evidence shows organisations should do and what most organisations are actually doing is the governance challenge this whitepaper has tried to map.

Figure 3. The AI Adoption-Transformation Gap



Notes: Organisational AI adoption is near-universal but workflow and operating-model redesign in response to it is rare. Cited surveys differ in methodology and definition, but the directional gap between adoption and substantive transformation is consistent across all major sources.

SECTION 3

The Pyramid under Pressure: Building Future Leaders

Of all the structural consequences of AI, the one most under-weighted in current decision-making is the effect on early-career development and the long-run leadership pipeline. That is partly because the consequence reveals itself slowly. The cost of undermining the pipeline does not appear on this year's profit and loss account, and in Australia, where many large organisations run lean graduate cohorts and rely on offshore professional services for entry-level work, the compression is already visible.

The entry-level compression problem

Most large organisations have historically operated a developmental model that resembles a pyramid. Junior employees perform the highest volume of routine cognitive work, first drafts, basic analysis, data compilation, initial code, and through repetition they build the pattern recognition, domain knowledge, and professional judgement required to move into more senior roles. That model is under structural pressure because AI now performs many of the tasks that previously constituted the entry-level curriculum.

The Daniotti study referenced in Section 2 is particularly relevant here. Across 160,000 developers in six countries, early-career developers were the most frequent users of AI coding tools, yet they derived no statistically significant productivity gains and showed no measurable expansion in technical breadth.

Senior developers using the same tools showed substantial improvements on both dimensions. Rather than closing the experience gap between junior and senior developers, unguided AI adoption appears to be widening it (Daniotti et al., 2026). Read alongside the Cui et al. randomised controlled trials in the same context, the message is clear: the developmental outcome depends on design, not on access.

Shukla and Hauge (2026) describe this from the practitioner side: as routine work moves into AI systems, the "learning laboratory that created the next generation of senior professionals" goes with it. Their proposed response - positioning juniors as AI supervisors rather than as AI replacements, paired with deliberate rotation and structured debriefs converges with the design principle implied by the empirical evidence.

Some executives have responded by reducing graduate intake or narrowing the scope of junior roles to the tasks AI cannot yet perform. The response is understandable in the short term. In the medium term, it may be strategically self-defeating. Thin the base of the pyramid today and you thin the pool of experienced, capable leaders available in ten years. The logic is straightforward, but the time horizon makes it easy to defer.

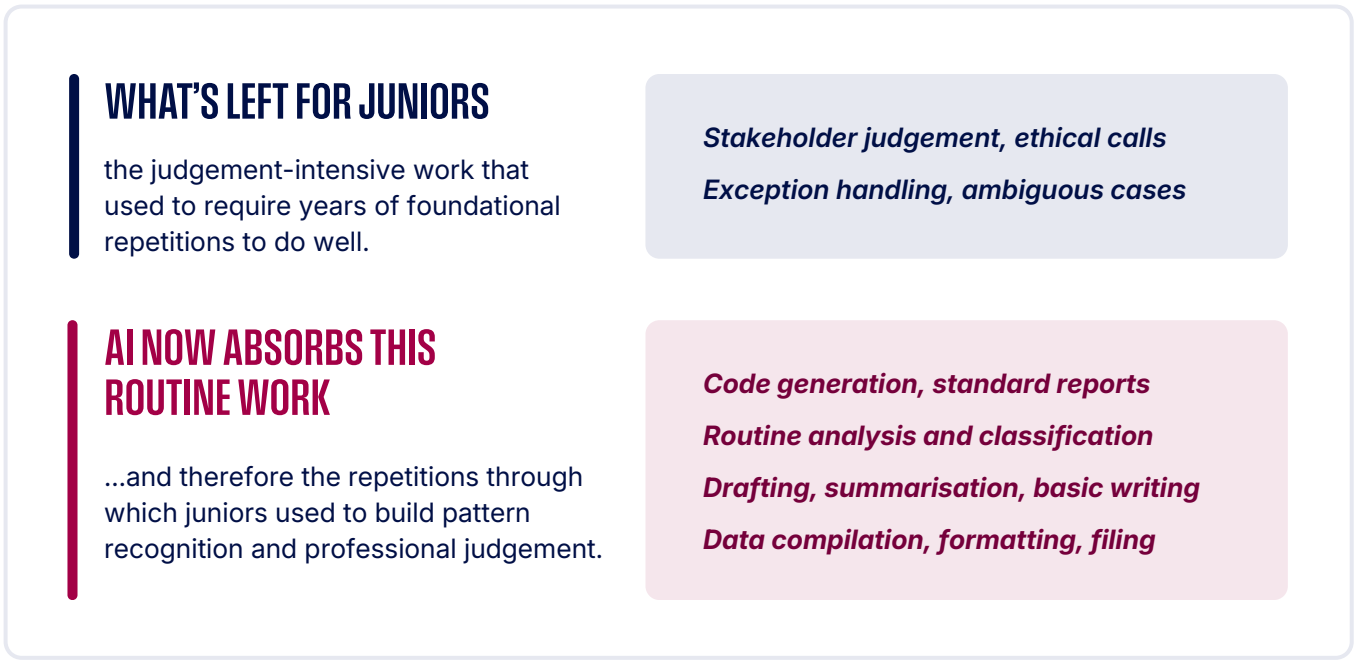


Figure 4. The Disappearing Curriculum: Hypothetical Junior Workload Before and After AI.
AI has compressed the apprenticeship: the routine tasks that once trained juniors into seniors no longer sit between them and the judgement-heavy work.

Using high-frequency payroll records from ADP, the largest US payroll processor, Brynjolfsson, Chandar, and Chen (2025) track employment by age and occupational AI exposure across millions of workers. They find that since generative AI's broad diffusion in late 2022, workers aged 22 to 25 in the most AI-exposed occupations have experienced a relative employment decline of around 13%, even after controlling for firm-level shocks, while employment for older workers in the same occupations and in less-exposed occupations has remained stable or grown. The declines appear in headcount rather than pay and are concentrated where AI automates rather than augments work, which is exactly why slower-moving aggregate statistics are late to register an entry-level hollowing-out (Brynjolfsson et al., 2025). KPMG's Q4 2025 AI Pulse Survey finds that 64% of organisations have altered their approach to entry-level hiring due to AI agents, up from 18% the previous quarter (KPMG, 2026). Together, these data suggest that the strategic question now facing executives is whether their organisation has a credible alternative pathway for developing the leaders they will need in 2030 and beyond.

There is a second, subtler risk. When early-career employees are given access to AI tools without the repetitions that build foundational competence, organisations may be creating what we would call human-capital debt. Work is produced, but the worker cannot explain, challenge, or improve upon it. Over time, that could weaken institutional memory, reduce resilience to error, and degrade the organisation's capacity for independent judgement when AI systems fail or mislead.

This pattern has now been observed experimentally. Bastani et al. (2025), in a randomised trial with nearly a thousand high school mathematics students, found that access to a standard ChatGPT-style tutor raised practice grades by 48% but lowered subsequent unassisted exam grades by 17% against controls. But a tutor designed with pedagogical guardrails eliminated the loss. Shen and Tamkin (2026) document the same pattern among developers learning a new programming library. Measured productivity with AI masked a degradation in conceptual understanding, code reading, and debugging. Performance with the tool can reflect either genuine capability or borrowed competence that disappears once the tool is withdrawn. Gillespie et al. (2025) report that more than three-quarters of students globally have

felt unable to complete their work without AI, and four in five put less effort into their studies and assessment knowing they can rely on it. Their findings suggest that the human-capital debt problem is already being incurred before workers enter the labour market.

The evidence for the human-capital debt problem is not confined to only students or juniors. Dell'Acqua et al. (2026), in a field experiment with 758 BCG consultants, find that on a task outside the AI's capability frontier, consultants using GPT-4 were 19 percentage points less likely to produce a correct recommendation. Yet their answers were rated 18–25% more coherent and persuasive - whether right or wrong. Even highly skilled professionals could not reliably tell when they had crossed the frontier. Human-capital debt shows up not only as missing competence in juniors, but as confidently wrong work in experienced staff.

A new design brief for talent development

The appropriate response is not to withhold AI tools from junior employees to preserve developmental difficulty. That approach is paternalistic, strategically naive, and does not address the underlying design problem.

The design problem is this. Entry-level roles need to be reconceived from first principles. They cannot be defined residually as whatever work remains after automation. They need to be designed affirmatively around a single question: what capabilities does this organisation need its future leaders to have developed by the time they reach mid-level positions, and what structured experiences will build those capabilities in an environment where AI does the automatable work?

Recent evidence offers a partial answer, albeit from the world of students in schools. Chung et al. (2026), in a five-month randomised trial across ten Taipei high schools, show that a GenAI chatbot *tightly integrated with adaptive problem-sequencing* improved exam performance by 0.15 standard deviations by sustaining engagement and productive difficulty. This size of effect would otherwise require six to nine months of additional schooling. The developmental architecture wrapped around the use of AI can significantly enhance learning.

This reframing shifts the challenge onto middle management. Managers in an AI-augmented environment will increasingly become coaches and validators rather than allocators of routine tasks. Their role is to design developmental experiences, to create the conditions in which juniors must exercise judgement rather than produce output, and to model the evaluative disposition that AI cannot replicate. Many current managers may not have been developed for this role, so organisations should invest in retraining them for it.

The emerging practice of agentic AI deployment offers an instructive parallel. At Ema, an agentic AI platform provider, there is no recruiter in the traditional sense. A "mission owner" for productive hires is accountable for the full journey, including sourcing, IT setup, compliance, training, team introductions, and early performance reviews, until a new hire is fully contributing. Success is measured not by tasks completed but by a crisp outcome KPI, "time to first commit," or how quickly software engineers deliver real code (Mantia, Chatterjee, and Lee, 2025). The model illustrates how organisations can redesign developmental architecture around outcomes rather than activity, a principle equally applicable to the challenge of early-career formation.

> ***Juniors will have AI whether employers provide it or not. Discussions about whether they should have AI are moot. The important question is what developmental architecture will build the leaders this organisation needs in 2030 and beyond.***

SECTION 4

The Human Premium: What People Are For

As AI takes on more of the cognitive work that humans have historically performed - drafting, summarising, coding, pattern-matching, and classifying - it becomes both more urgent and more tractable to articulate what humans are distinctively for. That is a clear strategic necessity. As Shukla and Hauge (2026) put it, the Industrial Age taught us to celebrate human consistency, speed, and rule following precisely because machines could not yet deliver them; AI now reveals that this was never really the source of human value.

The upward migration of human value

The research literature points in a consistent direction. As AI assumes more routine and pattern-based cognitive work, the relative value of human contribution migrates toward a cluster of capabilities that AI currently cannot replicate and may not replicate within any planning horizon relevant to today's institutions. Those capabilities cluster around four domains.

The first is judgement under genuine uncertainty. The capacity to make consequential decisions when the data are incomplete, the model is unreliable, or the ethical stakes are contested. AI can generate options and assess probabilities. It cannot bear accountability for choices, making judgement skill more valuable than ever before. But it also means that humans exercising judgement must be aware of the human biases they carry. In other words, judgements are not value-neutral, and therefore organisations will likely need to invest in developing better judgement that mitigates the biases we hold as human beings.

A second capability is relational and ethical intelligence. The capacity to read stakeholders, navigate political complexity, build trust, exercise moral reasoning, and take responsibility for outcomes that affect people. Far from being merely 'soft', these capabilities are the foundations on which organisational legitimacy rests.

The third capability is creative synthesis across domains. The capacity to frame problems in novel ways and combine insights from disparate fields rather than producing optimised recombinations of existing patterns. AI excels at the latter; the former remains distinctively human.

Fourth, there is institutional memory and contextual wisdom. The capacity to know what has been tried before, why it failed, what informal constraints govern this organisation, and how this decision will play out in six months. This knowledge is relational and largely irreducible to data.

AI is strongest at...	Humans remain essential for...
Rapid processing of large volumes of text, code, and pattern-based analysis	Defining the problem and the acceptable standard of judgement
Generating drafts, options, summaries, and likely next steps at speed	Choosing what matters, what is credible, and what should not be done at all
Consistency within known rules, templates, and bounded workflows	Ethics, accountability, empathy, and the navigation of stakeholder trade-offs
Scaling technical logic and surfacing patterns across large data sets	Creativity under uncertainty, political judgement, and social legitimacy
Orchestrating multi-step workflows across fragmented systems (agentic AI)	Stewardship: defining values, setting guardrails, deciding when to override automated decisions

The danger of outsourced thinking

The same capabilities that make AI a powerful augmentation tool create a structural temptation: to delegate not just tasks, but thought itself. The temptation is most acute among early-career employees, who may accept AI output without developing the evaluative capacity to challenge it. It is not absent in more experienced workers or even senior leaders.

Research on human-AI interaction identifies several cognitive risks from over-reliance on automated systems. Automation bias, where users accept machine outputs despite contrary evidence, has been documented across domains from aviation to medical diagnosis (Parasuraman and Riley, 1997). Skill atrophy from repeated delegation to AI systems can erode workers' capacity to perform tasks independently (Carr, 2010). Additionally, when individuals outsource monitoring functions to AI, they may engage in what cognitive scientists call cognitive offloading, reducing their oversight of decision quality (Risko and Gilbert, 2016). A global study of 48,000 people across 47 countries by Gillespie et al. (2025) found that two in three employees rely on AI output without evaluating it, and over half report having made mistakes in their work because of AI. These findings make the failure modes described previously measurable rather than theoretical.

Shaw and Nave (2026) measure the failure mode directly. Across three preregistered experiments with 1,372 participants, they show that when AI is available people consult it on most trials and adopt its answers even when those answers are systematically wrong. Accuracy rose 25 percentage points when the AI was correct and fell 15 percentage points when it erred, with confidence remaining elevated either way. They label this *cognitive surrender* and distinguish it from cognitive offloading (strategic delegation): a tipping point at which reliance on AI displaces rather than supplements human judgement.

These failure modes are not hypothetical or observed only in academic experimental studies. They are also visible in early clinical and legal deployments of AI tools, where practitioners have signed off on AI-generated outputs that contained material errors, sometimes because the volume and fluency of the output overwhelmed their evaluative capacity. The institutional response needs to be deliberate: *building evaluation skills* alongside deployment, *measuring decision quality* rather than output speed, and *creating cultural norms* in which questioning AI output is a mark of professional rigour, not technological scepticism.

With AI gaining a reputation for speed ("we can do everything so much faster now"), there is a significant risk that humans will now also be expected to do things faster. A natural reaction to such expectations, even demands, is that humans start outsourcing the hard work - thinking - to AI. But the risks are noted above. Organisations should therefore consider carefully their objective function. If it is speed, then outsourcing is almost inevitable. If it is better outcomes that overweight effectiveness, then some tasks of that work could be done faster but other tasks might just take longer than before. That should not only be acceptable but also expected by organisations.

Stewardship as an emerging leadership capability

The shift from tasks to outcomes, which is most pronounced in agentic AI deployments, introduces a fifth domain of human value that deserves explicit recognition - stewardship. As agentic AI systems gain the capacity to reason and act autonomously across departments and workflows, the human role shifts from directing individual tasks to stewarding intelligent systems. Leaders need to be able to encode organisational values into agent behaviour, define escalation paths, embed policy constraints, and maintain real-time monitoring with auditable logic and traceable logs (Mantia, Chatterjee, and Lee, 2025).

Stewardship is different from management in the traditional sense. It involves growing comfortable delegating to probabilistic systems one cannot fully script, while maintaining the governance structures that ensure every automated workflow reflects leadership values: what outcomes it prioritises, how it handles exceptions, and whose needs it serves. Organisations that cannot develop this capability will find themselves unable to govern the systems on which their operations increasingly depend.

SECTION 5

From Efficiency to Substance: The Redesign Imperative

Sections 3 and 4 argued that AI compresses the foundations of capability-building and challenges our view of what humans are distinctively for. This section turns to the operating model. Three shifts, closely related, need to be addressed together: the concentration of cognitive load, the widening gap between the appearance and the substance of productivity, and the change in what managers are being asked to do.

> *Capacity released by AI should not automatically be filled by more work. If every efficiency gain is converted into additional demand, organisations may find themselves with high output metrics and a deteriorating quality of human judgement.*

Concentrated cognitive load

By automating low-complexity work, AI may make the human experience of work harder rather than easier. If routine cognitive tasks are increasingly performed by machines, the work left to humans is more likely to involve exception handling, moral complexity, interpersonal conflict, creative ambiguity, and high-stakes accountability. Work becomes less repetitive but more intense and arguably more difficult and taxing.

Nobel Prize winner Daniel Kahneman's framework on System 1 and System 2 thinking is directly relevant here. System 1 processing is fast, automatic, and low-effort. System 2 processing is slow, deliberate, and high-effort (Kahneman, 2011). Much of what AI is removing from human work falls into System 1 territory: routine pattern-matching and first-pass drafting. What remains is disproportionately System 2, the cognitively demanding, emotionally taxing, and accountability-laden decisions that machines cannot reliably make.

An organisation that removes System 1 work from its human workforce and then fills the released capacity with additional System 2 demands has concentrated cognitive load rather than reduced it. If the productivity dividend of AI is systematically reinvested into higher throughput rather than human recovery, the result may be a workforce that is technically more productive in the short term and cognitively depleted in the medium term.

The performance theatre problem

A related risk is the widening gap between the appearance and the substance of productivity. As AI enables workers to produce high-volume, high-fluency output at speed, it also makes it easier to perform productivity without delivering it, generating reports, presentations, and communications that signal activity without creating commensurate value.

Chamorro-Premuzic (2026) provides the most incisive analysis of this dynamic. Work has always contained a performative dimension. Individuals are incentivised to manage impressions as much as to create value. The manager perpetually in meetings that produce no decisions, the employee generating polished reports that nobody reads, the executive curating a personal brand of action while delegating substance. These figures pre-date AI. But generative AI amplifies the phenomenon substantially. A report that once took a day to produce can now be assembled in an hour. A presentation that required a team can be created by one individual. The output exists but the thinking behind it may not.

Several principles help close the gap between appearance and substance. Focus on outputs that are as close as possible to value creation, even if imperfect, and resist the pull of easily measurable but low-value proxies. Triangulate performance using multiple sources of data, including peer and customer feedback, to counteract individual biases

and impression management. Create transparency around how work contributes to broader outcomes, making it harder to claim credit without substance. And update performance expectations in light of AI-driven efficiency gains, rewarding those who reinvest saved time into higher-value activities rather than those who simply maintain the status quo (Chamorro-Premuzic, 2026).

From directing to validating

The managerial role that AI demands is different in character from the one that many organisations have developed their leaders to perform. The traditional managerial function has emphasised directing and planning: allocating tasks, reviewing outputs, escalating exceptions, reporting upward. In an AI-augmented environment, much of the task allocation and first-pass output generation happens within automated or semi-automated workflows. The scarce managerial capability becomes something different: validation, probing, and assurance.

Validating means assessing whether AI-generated outputs are accurate, appropriate, and fit for purpose, a capability that requires genuine domain expertise, not merely process familiarity. Probing means asking whether the assumptions embedded in AI recommendations are sound, and whether the system is being asked to operate within or beyond its reliable frontier. Assurance, finally, is about maintaining accountability for the quality of decisions even when those decisions are informed, accelerated, or partially generated by AI. The shift from directing to validating has profound implications for leadership selection and development. The leaders who excelled in the previous model, high-throughput, decisive, and all about efficiency, may not be the leaders best suited to the new one. The new model favours evaluative depth and the capacity to identify and challenge the limitations of automated systems.

The agentic AI paradigm: designing around outcome

Agentic AI - systems capable of autonomous reasoning and action across fragmented applications and data sources - represents a shift from task automation to workflow orchestration. Rather than automating individual activities within existing processes, these systems can coordinate multi-step workflows across functions to achieve defined outcomes. In this agentic paradigm, AI systems reason and improve themselves.

As Mantia, Chatterjee, and Lee (2025) argue, the value of agentic AI is not realised through bolt-on efficiency gains, but through redesigning organisations around cross-functional execution rather than functional silos. This marks a qualitative shift from improving tasks within the current operating model to rethinking the operating model itself.

The practical implications are already visible. At Hitachi Digital, an agentic AI system named “Skye” serves as a first responder to employee HR queries across more than 20 disparate systems of record. Rather than requiring the organisation to consolidate its infrastructure into a single platform, the agentic system operates across existing, often siloed, systems, routing queries through an intent classifier to specialist agents for policy questions, document generation, leave management, IT help desk, and payroll. At NTT DATA, agentic AI was deployed to transform the process of responding to requests for proposal. AI agents gather information from various repositories within the company and the internet. A task that previously took weeks and required 20 people was replaced with AI agents that drafted complex 300-plus-page proposals in minutes. The human team refined sections collaboratively, but the AI agents handled the laborious groundwork. The combined human-AI effort improved efficiency more than three-fold (Mantia, Chatterjee, and Lee, 2025).

These examples illustrate a principle that extends well beyond individual deployments. Agentic AI forces operating model redesign. It is incompatible with functional silos. Agentic AI demands that organisations design around outcomes, not tasks. Every major journey needs what Mantia, Chatterjee, and Lee call a “mission owner,” someone who defines the mission, steers both humans and AI agents, and owns the outcome. That is a fundamentally different organisational model from functional silos in which marketing answers to brand, data answers to IT, and billing answers to finance. Customers do not experience companies this way.

They expect seamless, cross-functional resolution. Agentic AI makes the expectation technically achievable, but only if the organisation is redesigned to support it.

The economics of workforce redesign

Boards overseeing AI-driven workforce transitions are regularly presented with business cases that foreground gross labour savings. Those figures are typically more appealing than the full net economics, which need to include technology costs, token costs, integration costs, governance overhead, retraining investment, and transition costs (mostly cultural and social), all of which can materially alter the economic case.

Beyond the financial arithmetic, the structural consequences of aggressive headcount reduction deserve explicit scrutiny. Removing entry-level roles thins the leadership pipeline at a rate that is not immediately visible but will be felt in five to ten years. Large-scale displacement creates reputational consequences, for customer trust, employer brand, and employee morale, that may outweigh short-term savings. And automated workflows concentrate operational risk in ways that manual processes absorb. When AI systems degrade or produce errors at scale, the organisation needs the human expertise and institutional knowledge to identify and correct the failure quickly.

> *The primary board question is what operating model are we creating, and what risks are we importing into it that do not yet appear on the balance sheet?*

SECTION 6

The Governance Scenario: The LEDGER Test

To make these arguments concrete, consider a scenario familiar to many listed companies, including several that were represented at the roundtable. An organisation with a high labour-cost base and a credible AI deployment roadmap sees a path to substantially reducing its headcount within three years, primarily by automating entry-level and mid-level cognitive work across its professional services and operations functions. The financial model is compelling. The board is asked to approve an accelerated transformation programme.

Before approving, a well-governed board should put the proposal through what we call the LEDGER Test, a set of six questions designed to distinguish financially attractive proposals from strategically sound ones. A business case that cannot answer them with specificity and evidence is not ready for approval, regardless of the headline savings.

1 **L**

Licence

How are customers, regulators, employees, and the communities in which the organisation operates likely to interpret this decision? What is the reputational exposure, and how does it interact with the organisation's broader stakeholder commitments? In the Australian context, where large employers carry heightened scrutiny on domestic employment and where institutional investors increasingly integrate workforce transition into stewardship frameworks, this question has sharper teeth than it did a decade ago.

2 **E**

Economics

What are the net economics? What are the gross savings, and what are the costs of technology, integration, governance, retraining, and transition? How long is the payback period on the net figure, and how sensitive is that figure to delays in deployment or AI system performance shortfalls?

3 **D**

Dividend

If the programme delivers the projected efficiency gains, how will those gains be allocated? What share will be invested in capability formation, human sustainability, and the development of the next generation of organisational competence? What share will be returned to shareholders? These are legitimate governance questions, not rhetorical ones.

4 **G**

Governance of Agentic AI

If the transformation includes deployment of agentic AI systems that operate autonomously across departments and workflows, what governance architecture ensures that those systems encode leadership values, maintain auditable decision trails, and include clear escalation paths when outcomes diverge from intent? Who are the system stewards, and how is responsibility distributed between them, IT professionals managing security and data handling, and frontline teams empowered to intervene (Mantia, Chatterjee, and Lee, 2025)?

5 **E**

Endurance

What happens when the AI systems that replace these roles fail, produce errors at scale, or encounter novel situations outside their training distribution? Does the organisation retain sufficient human expertise to identify and correct these failures? How long would it take to rebuild that expertise if it is lost?

6 **R**

Renewal of Leadership and Capability Formation

If this programme proceeds, where will the organisation's future leaders come from in seven years? What is the current demographic of the senior leadership cohort, and what does the retirement pipeline look like? Is there a credible alternative development pathway that does not rely on the entry-level roles being eliminated?

SECTION 7

A Practical Agenda

The structural shifts described in this paper are either governed or not, and the consequences of non-governance accumulate invisibly until they become acute. What follows is a concrete agenda for the twelve months ahead, differentiated by institutional audience. Where a recommendation would duplicate a LEDGER Test question, we have left it to Section 6 and focused here on action rather than diagnosis.

For Chief Executives and Leadership Teams

The most consequential choices available to a CEO in the next twelve months are not about which AI tools to deploy. They are about what work, capabilities, and developmental pathways will exist on the other side of the deployment.

1. **Audit human-capital debt; do not merely track it.** Sample fifteen to twenty junior employees in your most AI-intensive functions. Ask each to explain or defend a recent piece of AI-assisted output, or improve on it. If fewer than half can do so adequately, the productivity gain is masking a capability problem. Repeat quarterly.

2. **Decide how the productivity dividend is allocated before it arrives.** Establish a board-approved policy on what proportion of expected efficiency gains goes to output growth, shareholder return, capability reinvestment, and human sustainability. Made retrospectively, this decision defaults to throughput. CFO and CPO co-own and report quarterly.

3. **Map the jagged frontier for your five most consequential roles.** Run a task-level analysis for each role across two axes: AI reliability and human consequence if AI is wrong. The high-consequence, low-reliability quadrant is your operational risk register. A week of structured workshops produces the map; functional leaders who cannot place their teams on it are not yet equipped to make sound deployment decisions.

4. **Change middle-manager performance criteria, not just their training.** The shift from directing to validating is real but remains unrewarded at this stage. Audit the metrics by which your top fifty middle managers are evaluated. Where those metrics still reward task throughput, rewrite them to include developmental challenge to direct reports and error-detection in AI-assisted workflows. For agentic AI deployments, appoint mission owners with cross-functional authority and outcome KPIs.

5. **Protect developmental difficulty by design.** The Daniotti finding - that unguided AI use widens the gap between junior and senior performance - is the empirical case for structure. Require first-draft attempts before AI is consulted on consequential tasks, structured critique of AI output as a deliverable, and verbal defence of reasoning. These are not restrictions on AI use. Rather, they are the repetitions through which professional judgement is built.

For Boards and Non-Executive Directors

The LEDGER Test in Section 6 provides the right diagnostic framework. Its limitation is institutional: a framework without a procedural home has no enforcement mechanism.

1. **Institutionalise the LEDGER Test as a standing approval condition.** Adopt a formal policy that any AI transformation programme exceeding a defined materiality threshold cannot be approved without a LEDGER-compliant brief. Assign the People and Culture committee primary responsibility for Renewal; assign Audit and Risk for Endurance and Governance of Agentic AI. Where agentic systems are deployed, require auditable decision logs and defined escalation paths.

2. **Require a capability age pyramid alongside every workforce business case.** Before approving headcount reduction, request the demographic profile of the senior leadership cohort, the retirement curve over five to ten years, and the pipeline of people who would historically have moved into those roles. If the entry-level base is being thinned and the senior cohort is skewing older, the organisation is trading short-term cost saving for a capability cliff in year seven. That is a board-level risk, not just an HR one.

3. **Commission an AI failure-mode scenario before each major deployment.** Ask management who in the organisation has the expertise to identify and correct errors if a specific automated workflow fails or degrades, and how long correction would take. Where the answer is uncertain, that is an operational risk requiring pricing. Review whether annual governance cadence matches the continuous pace of AI capability change.

For Policymakers

Aggregate employment statistics cannot detect the hollowing-out of entry-level work until it has happened. The current measurement architecture is structurally blind to the most consequential risk.

1. **Commission task-level labour market statistics.** Direct the ABS and Jobs and Skills Australia (JSA) to develop a task-composition supplement to occupational employment data: entry-level hiring rates by sector, the proportion of tasks now AI-assisted within high-exposure occupations, and time-to-promotion velocity for junior cohorts. The goal is a three-to-five year warning signal that current frameworks do not provide.

2. **Redirect transition policy toward early-career formation.** Most existing transition policy targets mid-career displacement. The harder structural risk is the removal of the first rung of the career ladder, which no re-skilling programme can later remedy. Establish a JSA workstream on entry-level pathway preservation, with co-investment incentives for employers that demonstrably redesign rather than eliminate junior roles.

3. **Create transparency obligations for large employers.** Modelled on gender pay gap reporting, require organisations above a defined threshold to disclose the ratio of AI-assisted to human-executed work in entry-level roles, graduate intake against prior years, and investment in early-career development. The mechanism is accountability through visibility, not regulatory mandate.

For Universities and Business Schools

Section 4 identifies the capabilities that constitute durable human value: judgement under genuine uncertainty, relational and ethical intelligence, creative synthesis, contextual wisdom, and stewardship. If these are what humans are now distinctively for, universities should be teaching them. Most are not, at least not explicitly. Tool familiarity is a hygiene factor; the deeper curriculum question is whether the foundations of professional judgement are being built deliberately or assumed.

1. **Teach the durable human capabilities as core curriculum, not as electives.** Most universities treat judgement under uncertainty, ethical reasoning, accountability, and stewardship as incidental byproducts of disciplinary learning, or relegate them to general education modules. Make them explicit graduate attributes with their own pedagogy and assessment, embedded in disciplinary contexts, for example diagnostic uncertainty in medicine, fiduciary accountability in finance, and evidentiary judgement in law. Case-based pedagogy that forces actual judgement, not knowledge recall, is the appropriate form. This requires faculty development as much as curriculum redesign.

2. **Replace submission-based assessment with interrogation-based assessment.** Provide students with AI-generated responses to discipline-specific problems. Assess their capacity to identify what is wrong, explain why, and produce a materially better answer. This tests the evaluative depth that constitutes durable human value, and is resistant to academic integrity concerns by design. Pilot in one programme per faculty in the coming academic year.

3. **Teach task-structure analysis as core professional literacy.** Decompose target occupations into task bundles and map them against AI susceptibility, augmentation potential, and durable human value. Use the Acemoglu-Autor framework and Dell'Acqua's jagged frontier as analytical tools. Each student produces a personal professional risk-and-opportunity map for their target role, revisited at graduation.

4. **Co-design the replacement entry-level curriculum with employers.** Convene a structured process with your highest-volume employer partners around one question: as graduate intake compresses, which developmental experiences are being eliminated, and how can a university employer partnership reconstruct them? The answer is more co-designed, outcome-measured developmental programmes. Model the governance question internally before advising others on it.

> *Together, these recommendations constitute a governance ecosystem: measurement instruments that make the test meaningful, data infrastructure that makes governance tractable, and institutional partnerships that reconstruct what the market and technology might be dismantling.*

SECTION 8

Signals to Watch: A Measurement Framework

Governance requires proper measurement. The following signals, tracked consistently over the next twelve months, will tell institutions whether AI-driven workforce transitions are proceeding in ways that are economically sound, developmentally sustainable, and socially legitimate.

Signals to track over the next twelve months

Signal	What to track	Why it matters
Entry-level hiring	Volumes and role mix by function; graduate intake versus AI tool deployment levels	Reveals whether the first rung of the ladder is being redesigned or removed
Wage effects	Wage pressure and trajectory in highly AI-exposed roles relative to control groups	Distinguishes productivity gains from wage devaluation in exposed segments
Output quality	Error rates, rework volumes, and decision quality in AI-assisted workflows	Tests whether speed gains reflect genuine value creation or deferred cost
Capability formation	Internal promotion rates, time-to-competence for junior cohorts, manager coaching load	Reveals whether the leadership pipeline is being strengthened or thinned
Social licence	Employee sentiment, employer brand, regulator attention, community response	Indicates whether the transition is sustainable beyond the current reporting period
AI adoption depth	Frequency of use, not just adoption rate; daily active use in core workflows	Distinguishes genuine integration from surface experimentation
Performance vs. theatre	Ratio of output volume to measurable outcome quality; customer satisfaction versus internal activity metrics	Detects whether AI is amplifying genuine productivity or its simulation
Agentic AI maturity	Proportion of workflows orchestrated end-to-end by agentic systems; resolution without human intervention; escalation frequency	Measures progression from task automation to outcome orchestration

CONCLUSION

Governing a Structural Transition

The evidence reviewed in this paper does not support either complacency or panic on the part of all stakeholders. It supports a more considered conclusion. AI is compressing some tasks, changing the mix of human capabilities that organisations need to develop and sustain, and forcing a structural reckoning with how work is designed, how talent is built, and how hierarchy functions. The emergence of agentic AI, systems that orchestrate entire workflows toward outcomes rather than automating individual tasks, intensifies these pressures while creating new organisational possibilities for those prepared to redesign around them. The changes are real, accelerating, and consequential, but they are also governable.

Organisations that treat AI as a narrow efficiency tool, deploying it to reduce headcount and reinvesting the savings in further throughput, may capture short-term gains while hollowing out the human foundations on which long-run performance depends. Those that use it to redesign work intelligently, to invest in human judgement, to protect the developmental pathways through which capability is built, and to maintain the social licence that legitimises their presence in the economies they serve, are more likely to emerge from this transition stronger.

This is, finally, not only a corporate governance question. In his first encyclical, *Magnifica Humanitas* (2026), signed on the 135th anniversary of Leo XIII's *Rerum Novarum*, Pope Leo XIV argued that "the pursuit of greater profits cannot justify choices that systematically sacrifice jobs," and that the economic order must remain subordinate to human dignity and the common good. The encyclical reframes the technology not as a problem to be forecast but as a choice about what kind of humanity its builders and deployers intend to serve - the same question, in a different register, that this paper has put to executives, boards, policymakers, and educators.

> ***The future of work is a design and governance question, and ultimately a question about what kinds of organisations and societies we are choosing to build. The window for deliberate action is open. The cost of waiting is not immediately visible, but it compounds.***

Authors

Prof. Ujwal Kayande

u.kayande@mbs.edu

Director - Institute for Digital Innovation and AI

Ujwal Kayande is Professor and Inaugural Director of the Institute for Digital Innovation & AI (IDIA) at Melbourne Business School. He previously founded the School's Centre for Business Analytics in 2014 and served as Director until 2021. He is a globally recognised expert in marketing analytics, AI, and digital transformation. Awarded several times for his research and teaching, he works closely with organisations to help them address the opportunities and challenges of our digital age. He is the Deputy Chair of the Board of Tech Futures Australia, which is Australia's leading digital technology education charity supporting 350,000+ school students nationwide. He is an Advisor to L'Oréal, Brandcomms.ai, T2 Tea, Lauretta.io, and Zetaris.

Kee Wong

kwong@e-ci.com.au

Non-Executive Director; Advisory Board Member, Institute for Digital Innovation and AI

Kee Wong is an accomplished Non-Executive Director with nearly four decades of senior executive and board experience across a diverse range of sectors. He currently serves as an independent Non-Executive Director for publicly listed companies, government enterprises, private businesses, and for-purpose organisations. A serial entrepreneur and global investor, Kee has founded and supported numerous start-ups and growth companies in technology, services, retail, food and beverage, and property. Kee served on the Board of the Australian Institute of Company Directors (AICD) for seven years from 2017 to 2023 and continues to Chair the AICD Governance of Innovation and Technology Panel. He presently holds Non-Executive Director roles with Car Group (formerly Carsales) globally, the Australian Energy Market Operator, the Australian Business Growth Fund, and Deloitte Australia.

The authors gratefully acknowledge the contributions of Professor Yalcin Akcay, Cheryl George (Director of Strategic Partnerships, IDIA), and Jacob McMahon (Coordinator, IDIA) to the development of this whitepaper. They also thank the participants of the IDIA roundtable for their generous time and input.

References

1. Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4, 1043–1171.

2. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *PNAS, 122(26)*, e2422633122.

3. Bessen, J. (2018). *AI and jobs: The role of demand* (NBER Working Paper No. 24235). National Bureau of Economic Research.

4. Bick, A., Blandin, A., & Deming, D. J. (2025). *The state of generative AI adoption in 2025*. Federal Reserve Bank of St. Louis. https://www.stlouisfed.org/on-the-economy/2025/nov/state-generative-ai-adoption-2025

5. Bloom, N., Sadun, R., & Van Reenen, J. (2012). Americans do IT better: US multinationals and the productivity miracle. *American Economic Review: Papers & Proceedings, 102(3)*, 167–171.

6. Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics, 117(1)*, 339–376.

7. Brynjolfsson, E., Chandar, B., & Chen, R. (2025). *Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence* [Working paper]. Stanford Digital Economy Lab. https://digitaleconomy.stanford.edu/publications/canaries-in-the-coal-mine/

8. Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work* (NBER Working Paper No. 31161). National Bureau of Economic Research.

9. Carr, N. (2010). *The shallows: What the internet is doing to our brains*. W. W. Norton & Company.

10. Chamorro-Premuzic, T. (2026, April 3). Pretending to work: AI, productivity, and the performative art of jobs. *Forbes*.

11. Chung, A., Zhang, B., Kung, L., Bastani, H., & Bastani, O. (2026). *Effective personalized AI tutors via LLM-guided reinforcement learning* (Wharton School research paper). SSRN. https://ssrn.com/abstract=6423358

12. Citadel Securities. (2026, February). *The 2026 global intelligence crisis* [Macro strategy note by F. Flight].

13. Citrini, & Shah, A. (2026, February 22). *The 2028 global intelligence crisis: A thought exercise in financial history, from the future*. Citrini Research.

14. Comyn, M. (2026, May 25). Pretending AI doesn't mean change won't protect workers. *Australian Financial Review*.

15. Cui, K. Z., Demirer, M., Jaffe, S., Musolff, L., Peng, S., & Salz, T. (2025). *The effects of generative AI on high-skilled work: Evidence from three field experiments with software developers* [Working paper].

16. Daniotti, S., Wachs, J., Feng, X., & Neffke, F. (2026). Who is using AI to code? Global diffusion and impact of generative AI. *Science, 391*(6787), 831–835.

17. Dell'Acqua, F., McFowland, E., III, Mollick, E., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science*. Advance online publication.

18. Deloitte. (2026). *State of AI in the enterprise: 2026 AI report*. Deloitte AI Institute.

19. Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: An early look at the labor market impact potential of large language models. *Science, 384*(6702).

20. Felten, E. W., Raj, M., & Seamans, R. (2019). *The occupational impact of artificial intelligence: Labor, skills, and polarization*. NYU Stern School of Business.

21. Felten, E. W., Raj, M., & Seamans, R. (2023). *How will language models like ChatGPT affect occupations and industries?* (NBER Working Paper No. 31053). National Bureau of Economic Research.

22. Floridi, L., Novelli, C., & Morley, J. (2026). *Not even wrong: Forecasts of AI impact on the labour market, from Frey-Osborne to ChatGPT, 2013–2026.* Digital Ethics Center, Yale University. SSRN. https://ssrn.com/abstract=6830981

23. Gillespie, N., Lockey, S., Ward, T., Macdade, A., & Hassed, G. (2025). *Trust, attitudes and use of artificial intelligence: A global study 2025*. The University of Melbourne and KPMG.

24. Girotra, K., Meincke, L., Terwiesch, C., & Ulrich, K. (2023). *Ideas are dimes a dozen: Large language models for idea generation in innovation* [Working paper]. SSRN.

25. Gmyrek, P., Berg, J., & Bescond, D. (2023). *Generative AI and jobs: A global analysis of potential effects on job quantity and quality* (ILO Working Paper No. 96). International Labour Organization.

26. Goldman Sachs. (2025, August). *How will AI affect the global workforce?* Goldman Sachs Research.

27. Hartley, J. S., Jolevski, F., Melo, V., & Moore, B. (2025). *The labor market effects of generative artificial intelligence* [Working paper, revised September 2025]. SSRN. https://ssrn.com/abstract=5136877

28. Johnston, A. C., & Makridis, C. A. (2026). *AI, output, and employment* [Working paper]. SSRN. https://ssrn.com/abstract=6460358

29. Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

30. KPMG. (2026, January 15). *AI quarterly pulse survey, Q4 2025: AI at scale — How 2025 set the stage for agent-driven enterprise reinvention in 2026*. KPMG US.

31. Leo XIV. (2026). *Magnifica humanitas: On safeguarding the human person in the time of artificial intelligence* [Encyclical letter]. The Holy See.

32. Mäkelä, P., & Stephany, F. (2024). *Complement or substitute? Occupational exposure and AI adoption patterns* [Working paper]. Oxford Internet Institute.

33. Makridis, C. A. (2026, April 14). Industries most exposed to AI are not only seeing productivity gains but jobs and wage growth too. *The Conversation*.

34. Makridis, C. A., & Brynjolfsson, E. (2026). *Counting AI: A blueprint to integrate AI investment and use data into U.S. national statistics*. Brookings Institution.

35. Mantia, L., Chatterjee, S., & Lee, V. S. (2025, October 24). Designing a successful agentic AI system. *Harvard Business Review*.

36. Massenkoff, M., & McCrory, P. (2026, March). *Labor market impacts of AI: A new measure and early evidence*. Anthropic Research.

37. McKinsey & Company. (2025, November). *The state of AI global survey 2025: How organizations are rewiring to capture value*. McKinsey QuantumBlack.

38. Meyersohn, N. (2026, April). The demise of software engineering jobs has been greatly exaggerated. *CNN Business*.

39. Morgan Stanley. (2026). *AI and jobs: Labor market impact echoes past tech transitions.* Morgan Stanley Research.

40. Munis, J. (2026, February 17). Klarna's CEO agrees with Dario Amodei. He thinks his white-collar workforce will shrink by a third by 2030. *Fortune*.

41. Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative AI. *Science, 381*(6654), 187–192.

42. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253.

43. PwC. (2025, May). *PwC's AI agent survey*. PricewaterhouseCoopers.

44. Quiroz-Gutierrez, M. (2026, February 25). Jamie Dimon says society should start preparing for AI job displacement: "Now's the time to start thinking about" it. *Fortune*.

45. Ransbotham, S., Khodabandeh, S., Kiron, D., LaFountain, B., & Candelon, F. (2025, November 18). *The emerging agentic enterprise: How leaders must navigate a new age of AI*. MIT Sloan Management Review and Boston Consulting Group.

46. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688.

47. Shaw, S. D., & Nave, G. (2026). *Thinking — fast, slow, and artificial: How AI is reshaping human reasoning and the rise of cognitive surrender* (SSRN Working Paper No. 6097646). SSRN.

48. Shen, J. H., & Tamkin, A. (2026). *How AI impacts skill formation* [Working paper]. arXiv. https://doi.org/10.48550/arXiv.2601.20245

49. Shukla, M., & Hauge, N. (2026). *The five-year century: Bold leadership and accelerated outcomes in the age of AI*. John Wiley & Sons.

50. Stanford Institute for Human-Centered Artificial Intelligence. (2026). *The 2026 AI index report*. Stanford University. https://hai.stanford.edu/ai-index/2026-ai-index-report

51. The Budget Lab at Yale. (2026, April 16). *Tracking the impact of AI on the labor market*. Yale University.

52. U.S. Bureau of Labor Statistics. (2025). *Occupational outlook handbook: Software developers*. U.S. Department of Labor. bls.gov.

53. U.S. Bureau of Labor Statistics. (2026, March 31). *Job openings and labor turnover summary (JOLTS)*.

54. VandeHei, J., & Allen, M. (2025, May 28). Behind the curtain: A white-collar bloodbath. *Axios*.

55. Weber, L. (2026, February 26). What AI executives tell their own kids about the jobs of the future. *The Wall Street Journal*.

Institute for DIGITAL INNOVATION & AI

The Institute for Digital Innovation and AI (IDIA) at Melbourne Business School helps organisations understand and realise the opportunities of AI and digital technologies. We combine academic insight with industry experience to deliver practical skills that turn AI potential into business impact.

IDIA partners access resources, events, and insights to identify opportunities, upskill teams, and implement responsible, scalable transformation. Contact idia@mbs.edu to learn more.

Connect with us

mbs.edu/idia
idia@mbs.edu

UJWAL KAYANDE
Professor and Inaugural Director of IDIA
u.kayande@mbs.edu

CHERYL GEORGE
Director, Strategic Partnerships
c.george@mbs.edu | +61 402 154 561

JACOB MCMAHON
Coordinator
j.mcmahon@mbs.edu

Melbourne Business School
THE UNIVERSITY OF MELBOURNE

Why Melbourne Business School?

We are uniquely positioned to deliver transformative learning experiences that align with your organisation's strategic goals.

WORLD CLASS EXPERTISE

MBS combines the academic excellence of Australia's top business school with the practical insight of experienced industry practitioners. Our programs are led by globally recognised faculty and expert associates who bring the latest thinking in AI, leadership, strategy, innovation, digital technologies, data analytics, and organisational development.

FLEXIBLE, MULTI-MODAL DELIVERY

From immersive face-to-face masterclasses to engaging virtual sessions and self-paced digital content, MBS offers flexible delivery options that suit your workforce's diverse needs and schedules.

TAILORED, CONTEXTUALISED SOLUTIONS

We don't offer one-size-fits-all programs. Every engagement is co-designed to reflect your organisation's unique challenges, context, and goals - ensuring relevance, buy-in, and immediate applicability.

COMMITMENT TO LONG-TERM PARTNERSHIPS

We collaborate closely with our clients to support long-term capability development - evolving programs over time to meet changing needs and sustain momentum.

Disclaimer

This whitepaper is intended to inform and spark discussion. It doesn't constitute legal, accounting, or other professional advice, and should not be relied on as a substitute for advice tailored to your circumstances. We have taken reasonable care in preparing it, but we cannot guarantee that the material is complete or error-free. The views expressed by the authors are their own. Any links to third-party sites are provided for convenience and do not imply endorsement.

Institute for
DIGITAL INNOVATION & AI

mbs.edu

Published in June 2026

THE UNIVERSITY OF MELBOURNE

Melbourne Business School