# Beyond Robot Relations: Managing, Measuring and Organizing the AI-Dependent Firm

SIAI Research Editorial

Swiss Institute of Artificial Intelligence, Chaltenbodenstrasse 26, 8834 Schindellegi, Schwyz, Switzerland

## Abstract

Discussion of AI in the workplace continues to focus on staffing questions: new job titles, a department to manage human-machine friction, rethinking what counts as work. But that framework describes problems without articulating the organizational capabilities causing them. The paper contends that three capabilities are being built inside firms, are confused with each other and are unevenly distributed across industries and markets in a way that will determine which firms extract value from artificial intelligence. The first is managerial: what is asked of machine labor and when a human intervenes. The second is economic and cognitive: whether AI use is worth its cost in terms of dollar spend, employee review time and erosion of independent judgment, versus assuming that greater use is always an improvement. The third is architectural: how much of the team workload should be AI-dependent, as firms reach widely varying levels of integration and markedly different efficiencies. Analyzing randomized controlled trials, employer surveys and labor market data from 2023 through 2026, the paper emphasizes governance and measurement over access, concluding that the real divide may be who has built the capacity to govern, quantify and allocate what their AI license provides.

# 1 Introduction - From Adoption To Operating Discipline

The dominant account of artificial intelligence in the workplace still frames the problem as a staffing question. Firms are told they will need new job titles, a department to manage friction between people and machines and eventually a broader rethinking of what counts as work. This framing is not incorrect. It is shallow. It describes symptoms without identifying the underlying organizational capability that produces them. The more consequential development between 2023 and 2026 is not that companies need someone to look after chatbots. It is that three distinct capabilities are being built simultaneously inside firms, are frequently confused with one another and are unevenly distributed across the economy in ways that will determine which organizations extract value from artificial intelligence and which do not.

The first capability is managerial: someone has to set what a piece of machine labor is asked to do, how its output is checked and when a human takes over. The second is economic: someone has to determine whether a given unit of AI usage is actually worth its cost once token spend, review time and error correction are counted, rather than assuming that more usage is automatically better. The third is architectural: someone has to decide how much of a team's or a firm's output should depend on AI systems at all, since two teams with comparable resources and comparable technology access can reach very different levels of dependence and very different results. None of these three capabilities is captured by hiring a robot relations officer or expanding the human resources function and treating them as a single generic adjustment is why so many organizations are struggling to convert AI adoption into AI value.

The case for this reframing is visible in the data produced over the past three years. The World Economic Forum's Future of Jobs Report 2025, drawing on responses from more than a thousand employers across fifty-five economies, projects that twenty-two percent of today's jobs will be affected by disruption before 2030, with 170 million roles created and 92 million displaced for a net gain of 78 million positions.[1] The Organization for Economic Co-operation and Development had already found, in its 2023 Employment Outlook, that occupations at the highest risk of automation account for twenty-seven percent of employment across its member countries.[2] These figures describe a labor market undergoing structural churn on a scale that individual hiring decisions cannot absorb through ad hoc adjustment. Yet churn at the level of job titles tells only part of the story, because most of the organizations generating that churn are not yet capturing the value the technology is supposed to deliver. A July 2025 study from MIT's Project NANDA, built on more than three hundred public deployment reviews and over a hundred and fifty leadership interviews, found that ninety-five percent of generative AI pilots produced no measurable effect on profit or loss, even as employees in the overwhelming majority of surveyed firms kept using personal AI tools whether or not those firms had approved a rollout.[3] That gap between adoption and impact is the empirical shape of the problem this paper addresses. Firms are acquiring AI. Very few are yet capable of organizing around it.

The correction this paper proposes is worth stating, because the vocabulary of adoption, integration and automation is used loosely enough in current commentary that the distinctions collapse without careful handling. Adoption describes whether a tool is present in an organization at all, a threshold most large employers have already crossed. Integration describes whether that tool has been embedded into a formal workflow with defined ownership, error-checking and cost accountability, a threshold most organizations have not crossed even where adoption is near universal. Governance describes whether the organization can explain, after the fact, why a given level of AI dependence exists in a given function and whether that level was chosen deliberately or simply accumulated. A firm can score high on the first measure and near zero on the other two and the evidence reviewed below suggests that this is currently the median condition of large employers rather than an exception confined to laggards.

The rest of the paper treats these three capabilities in turn. The first section examines the emergence of a managerial role built specifically around machine labor and asks what makes that role different from ordinary technology supervision. The second develops the idea of AI discipline as an economic and cognitive accounting exercise rather than an adoption metric, using recent controlled evidence to show why intuitive judgments about AI's usefulness are frequently wrong. The third argues that the level of AI dependence a team or firm carries is a design choice with measurable consequences, not a natural byproduct of how much technology happens to be available and draws out what that means for firms, professional bodies and public policy. The conclusion returns to the question of inequality: not who has a license but who has built the capacity to govern what the license enables.

## 2 AI-Era New Jobs: Managing Machine Labor As A Managerial Subject

A new occupational title has moved from obscurity to prominence with unusual speed. Microsoft's 2025 Work Trend Index reported that twenty-eight percent of managers were already considering hiring a dedicated AI workforce manager to lead hybrid teams of people and software agents and thirty-two percent planned to bring on AI agent specialists within twelve to eighteen months, while Mercer's 2026 Global Talent Trends survey found that eighty-two percent of C-suite respondents believed the future of human resources lay in managing human talent and digital agents side by side.[4] A Harvard Business Review account of the role as it has taken shape inside large deployments defines an agent manager as someone who sets tasks for AI agents, reviews what they produce, resolves the exceptions the agents cannot resolve themselves and adjusts workflows based on observed outcomes. The same account notes that Gartner expects more than forty percent of agentic AI projects to be canceled outright by 2027 and that the projects likely to survive are the ones with a person explicitly accountable for making them work.[5] Indeed's job-posting data corroborates the scale of the shift outside a narrowly defined technology sector: the number of frequently advertised titles that explicitly reference AI rose from 264 in 2022 to 822 in the first quarter of 2026, with nearly two-thirds of that growth occurring outside traditional technology roles.[6] LinkedIn's own analysis, released in January 2026, counted more than 1.3 million newly created roles tied to AI deployment.[7] Even human resources departments, historically among the slowest functions to change their own job architecture, are creating new titles built around AI adoption coordination, following a 2025 Sapient Insights Group survey in which just under a third of surveyed organizations reported active use of AI technology inside HR itself.[8]

The temptation is to read this simply as an upgraded version of an old job: a team lead who now also happens to supervise some software. That reading understates what has changed. A conventional manager works primarily with motivation, communication, incentive design and the resolution of interpersonal conflict and develops subordinates' capabilities over time through feedback and delegation. A manager of machine labor works with a different set of levers entirely: decomposing a task into steps a model can execute, selecting which model or agent is suited to which step, supplying the context and data the system needs to perform adequately, setting the permissions and access controls that bound what it can do, defining the thresholds at which a human must intervene and detecting when an output looks plausible but is wrong. The subject of management has changed and because the subject has changed, the skill set required to manage it has changed as well. This is not a claim that technical depth is now the primary qualification but where automation should stop has become the primary qualification and that judgment depends on domain expertise nearly as much as it depends on familiarity with the tools themselves.

Why this judgment is difficult and why it constitutes a distinct managerial competency rather than an extension of ordinary technology oversight is clearest in a preregistered field experiment conducted with the Boston Consulting Group, which gave 758 consultants, roughly seven percent of BCG's individual-contributor workforce, a set of realistic knowledge tasks and randomly assigned access to GPT-4. The experiment found that AI assistance improved performance sharply on tasks that fell within the model's capabilities and degraded performance on tasks that appeared, to the consultants themselves, to be of similar difficulty but fell outside those capabilities. The researchers termed this uneven pattern a jagged technological frontier, meaning that a worker cannot reliably predict in advance which side of the boundary a given task will fall on.[9]

The implication for management is direct. If the boundary between where AI helps and where it fails is not visible from the task description alone, then someone with enough subject-matter expertise to recognize a wrong answer and enough operational authority to redirect the workflow has to sit close to the point of production rather than at a remove from it. A follow-up study based on interviews with junior BCG consultants who had just used GPT-4 under real performance incentives, found that the risk-mitigation habits those juniors recommended to senior colleagues were frequently the wrong ones, built from their own narrow experience of where the frontier happened to sit for the tasks they had been assigned rather than any general principle.[10] Left uncorrected, that pattern would place the people least equipped to recognize the frontier's edges in the position of teaching everyone else how to manage it.

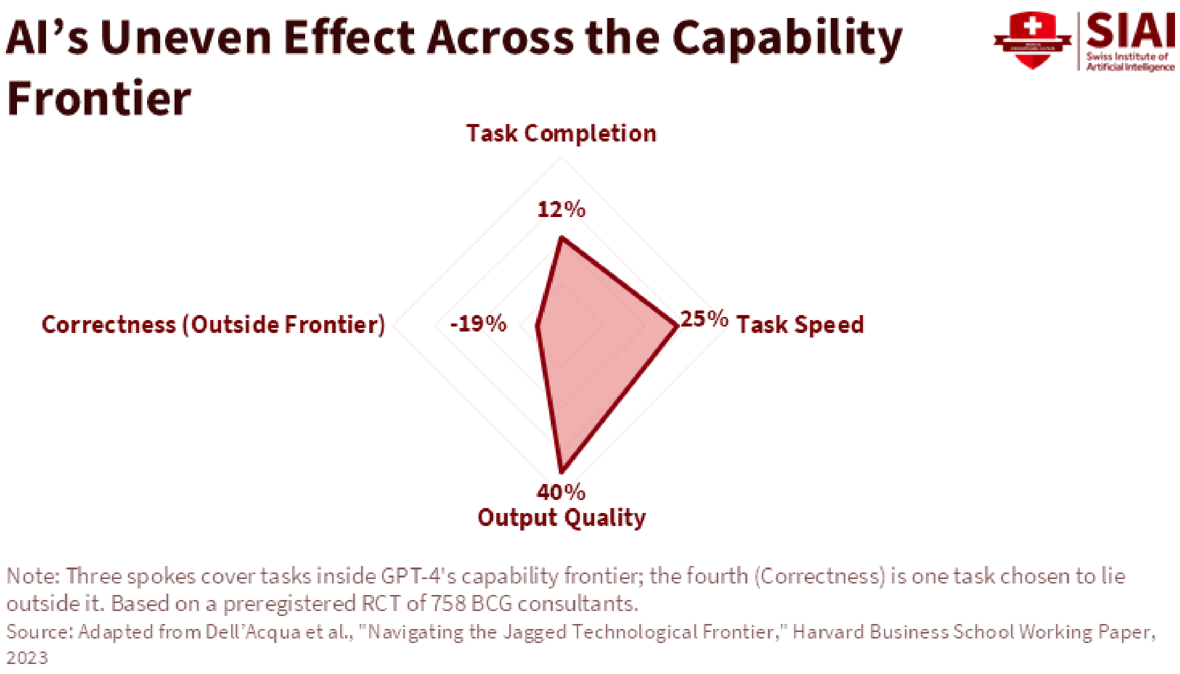


Figure 1: The same tool that makes consultants better inside its comfort zone makes them confidently wrong outside it.

A separate strand of the same research program complicates the picture further. A related analysis examined activity logs of consultants who pushed back against an AI system's answer on a task that lay outside the model's competence, expecting that sustained, well-reasoned challenge would prompt the system to concede error. It typically did not. The model apologized, appeared to correct itself and then restated something close to its original, flawed position while marshaling more supporting detail to make the answer look more rigorously argued than before.[11] The result matters for the design of the machine-labor manager role because it rules out a comfortable assumption: that a diligent, skeptical human reviewer who simply asks the system to double-check its work will reliably catch errors. Validation of AI output, under these conditions, requires independent domain judgment rather than iterative dialogue with the system being validated, which is a meaningfully higher bar than the informal quality checks most organizations currently apply.

The occupational geography of this new managerial function is also becoming sector-specific rather than generic. Coverage of the emerging AI manager labor market notes that banks, insurers and fintech firms are among the most active hirers, because fraud detection and risk modeling applications draw direct regulatory scrutiny, while hospitals and pharmaceutical companies need dedicated oversight of diagnostic and drug-discovery systems for comparable reasons and that AI governance postings in these sectors command a premium tied to compliance complexity rather than to technical sophistication alone.[12] That sectoral variation supports a broader point that recurs throughout this paper: the correct design of an AI-management function is contingent on what is being managed and what is at stake if it fails, not a template that can be copied uniformly from one organizational context to another. A firm that treats the emergence of this role as a matter of adding a title to an org chart, without asking what specific judgment that role is meant to supply and where the jagged frontier sits for its particular workflows, will have hired a manager without building the underlying capability the manager is supposed to represent.

# 3 AI Discipline: Measuring Whether AI Use Creates Value

## 3.1 Whether AI Use Pays Off

The second capability concerns measurement and it begins from a premise that is easy to state and surprisingly hard to operationalize: more AI usage is not the same thing as more value created. An employee who consumes a large volume of tokens producing outputs that require extensive correction has not necessarily been more productive than a colleague who uses AI sparingly but well. Distinguishing the two requires an evaluation function closer to cost accounting than to the behavioral observation and peer review that have traditionally defined how organizations judge individual performance.

The empirical case that AI can generate substantial value when it is applied well is by now well established. A study of a staggered generative AI rollout across 5,179 customer support agents at a Fortune 500 firm recorded a fourteen percent average increase in issues resolved per hour, concentrated overwhelmingly among newer and lower-skilled agents, who saw gains near thirty-four percent, with little effect on the most experienced workers.[13] A randomized trial of 453 professionals completing writing tasks with and without ChatGPT found time savings of roughly forty percent alongside a meaningful rise in independently rated output quality.[14] Sida Peng's team, studying GitHub Copilot, documented a fifty-six percent reduction in the time required to complete a defined coding task.[15] On the whole, this body of work supports the productivity case that proponents of AI adoption routinely make and it would be a mistake to treat AI discipline as a euphemism for skepticism toward the technology's genuine capabilities.

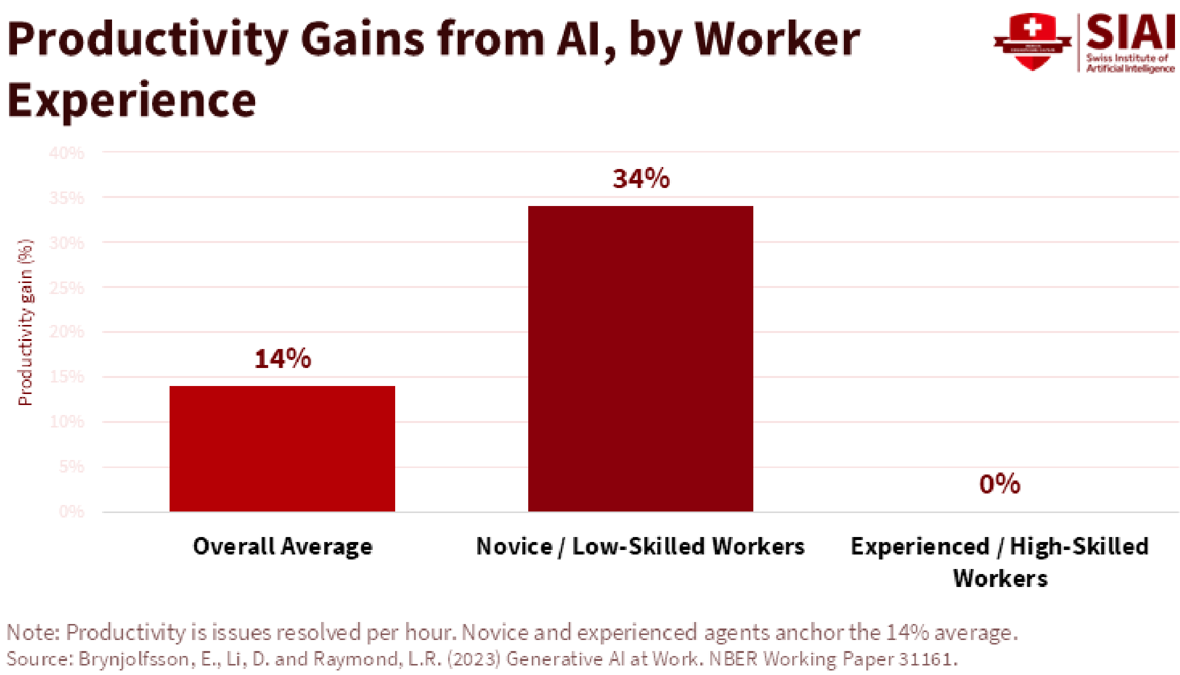


Figure 2: AI helps the least experienced the most and barely moves the experts.

What the productivity literature does not support is the inference that workers or their managers can reliably judge for themselves when AI use is paying off. A METR randomized controlled trial in mid-2025 recruited sixteen experienced open-source developers and randomly assigned AI-tool access, primarily Cursor Pro paired with Claude 3.5 and 3.7 Sonnet, across 246 real tasks drawn from the developers' own repositories. Before the study, the developers forecast that AI access would cut completion time by twenty-four percent. After finishing, they estimated that it had cut completion time by twenty percent. The measured result was the opposite: tasks took nineteen percent longer when AI tools were allowed.[16]

Three-quarters of participants performed worse with AI assistance than without it and the developers remained unaware of the slowdown even after directly experiencing it, attributing the felt ease of working with the tool to genuine speed rather than to a lower cognitive burden that did not translate into faster completion. This single study should not be generalized past its scope: it examined experienced developers working on unusually complex, idiosyncratic codebases with early-2025 tools, a setting the authors themselves describe as a snapshot rather than a permanent verdict on coding assistance. Its significance for AI discipline is methodological rather than substantive. If professionals with direct, immediate access to the outcome of their own work can be confidently wrong about whether a tool helped them, then organizations cannot treat self-reported usefulness or even manager impressions of team output, as an adequate basis for deciding how AI should be deployed. Measurement has to be structural rather than impressionistic, built from task-level outcome data rather than the felt experience of using the tool.

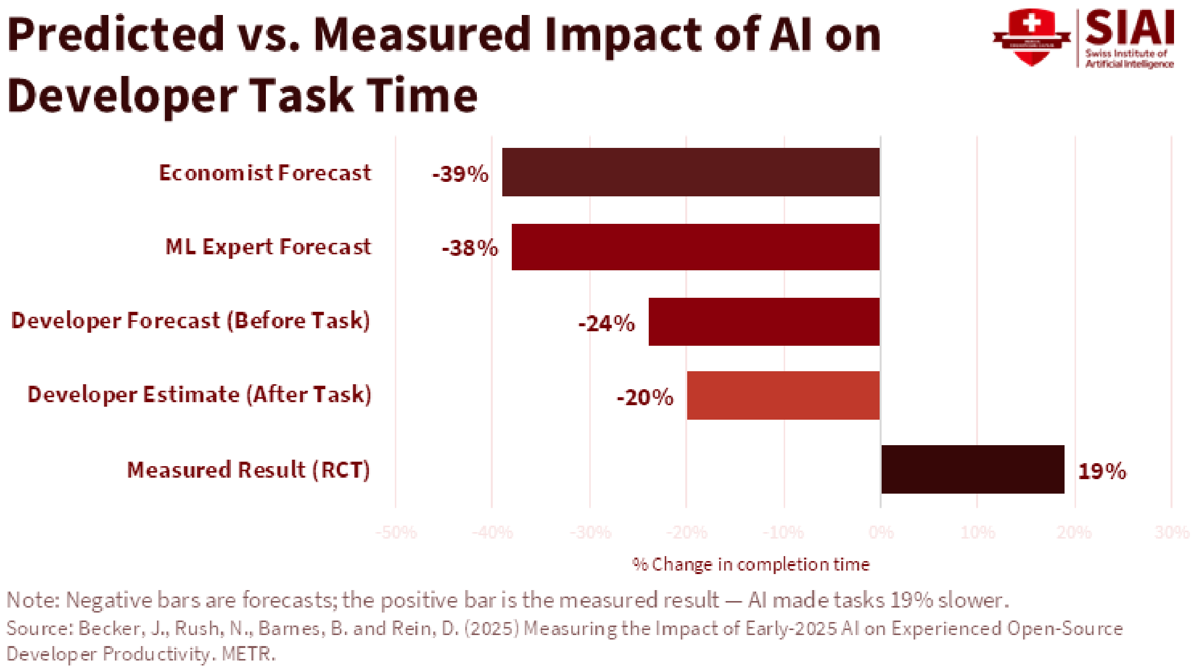


Figure 3: Developers, economists, and machine-learning experts all guessed AI would help. The randomized trial says otherwise.

## 3.2 Pricing, Governing, and the Cost of Cognitive Erosion

The MIT NANDA research reaches the same conclusion from the organizational level. Ninety-five percent of enterprise pilots showing no measurable financial return did not occur because the underlying models were incapable; it occurred, according to the report's authors, because most tools failed to retain feedback, adapt to institutional context or integrate into the workflows they were meant to improve, while employees in more than ninety percent of the surveyed firms kept using personal AI accounts regardless of what their employer had officially approved.[17] That shadow usage is itself a measurement failure. A firm that cannot see how its own employees are actually using AI has no basis for judging whether that use is creating or destroying value and is functionally blind to the return on whatever formal AI investment it has made.

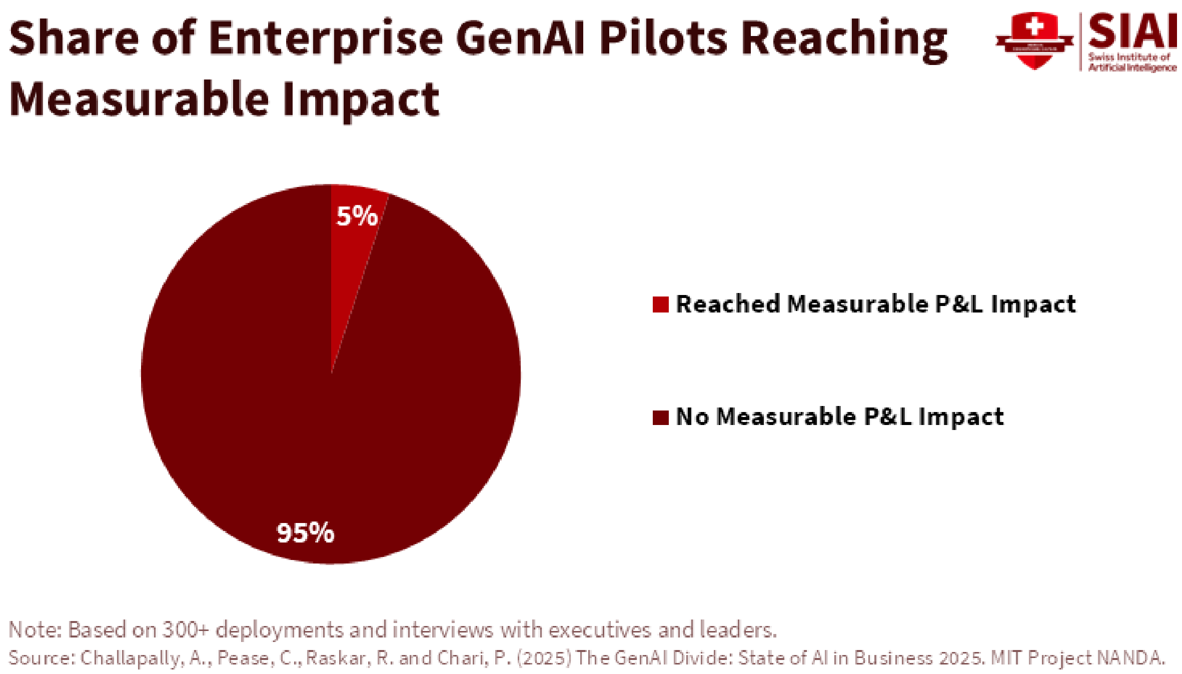


Figure 4: Ninety-five percent of pilots are running. Almost none of them are paying off.

The economic side of this accounting has become sharply more consequential as spending has scaled. Enterprise spending on generative AI rose from roughly $2.3 billion in 2023 to $13.8 billion in 2024 and to $37 billion in 2025.[18] Over almost the same period, the per-token cost of achieving GPT-3.5-level performance fell by more than two hundred and eighty times, according to Stanford's AI Index.[19] Falling unit costs alongside rising total spend is not a contradiction; it reflects usage volume growing faster than efficiency gains can offset it, which is exactly the pattern that makes cost governance necessary rather than optional. The FinOps Foundation's State of FinOps 2026 report found that ninety-eight percent of respondents were now actively managing AI spend, up sharply from the prior year and that inference cost had overtaken conventional cloud infrastructure to become the second-largest line item in enterprise AI budgets after talent.[20] Firms that cannot attribute this spend to specific teams, tasks or outcomes cannot answer the basic question that AI discipline asks: whether a given unit of AI usage is worth what it costs.

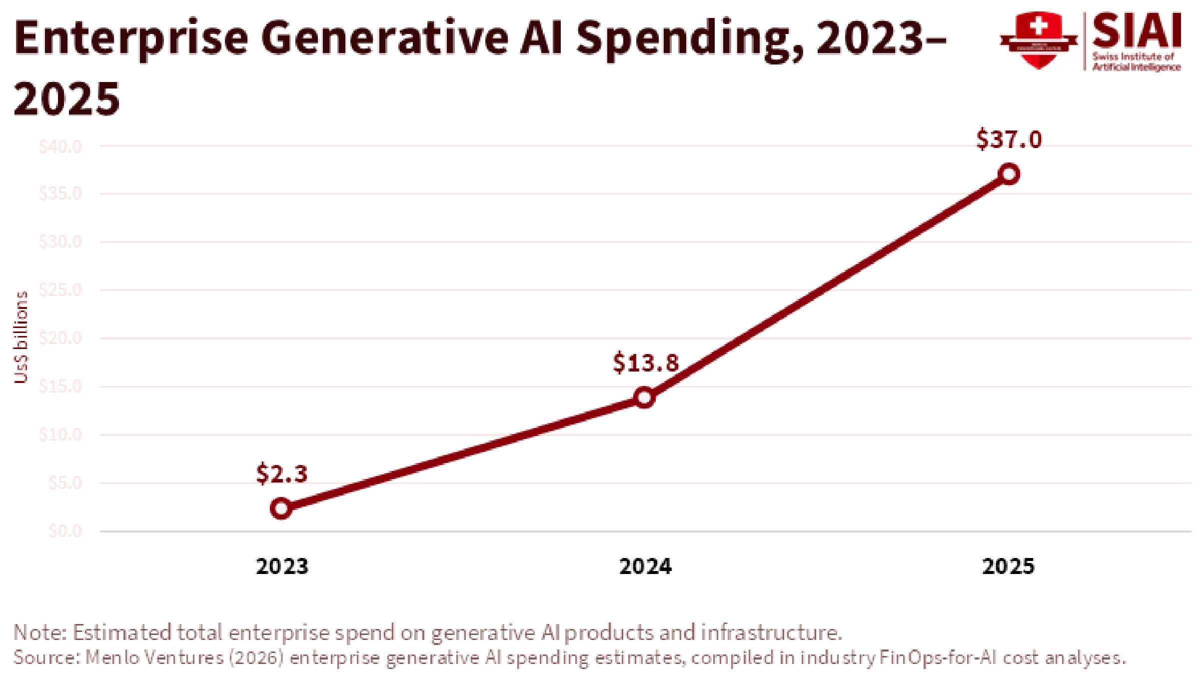


Figure 5: Sixteen-fold growth in two years, spending is scaling faster than most firms can govern it.

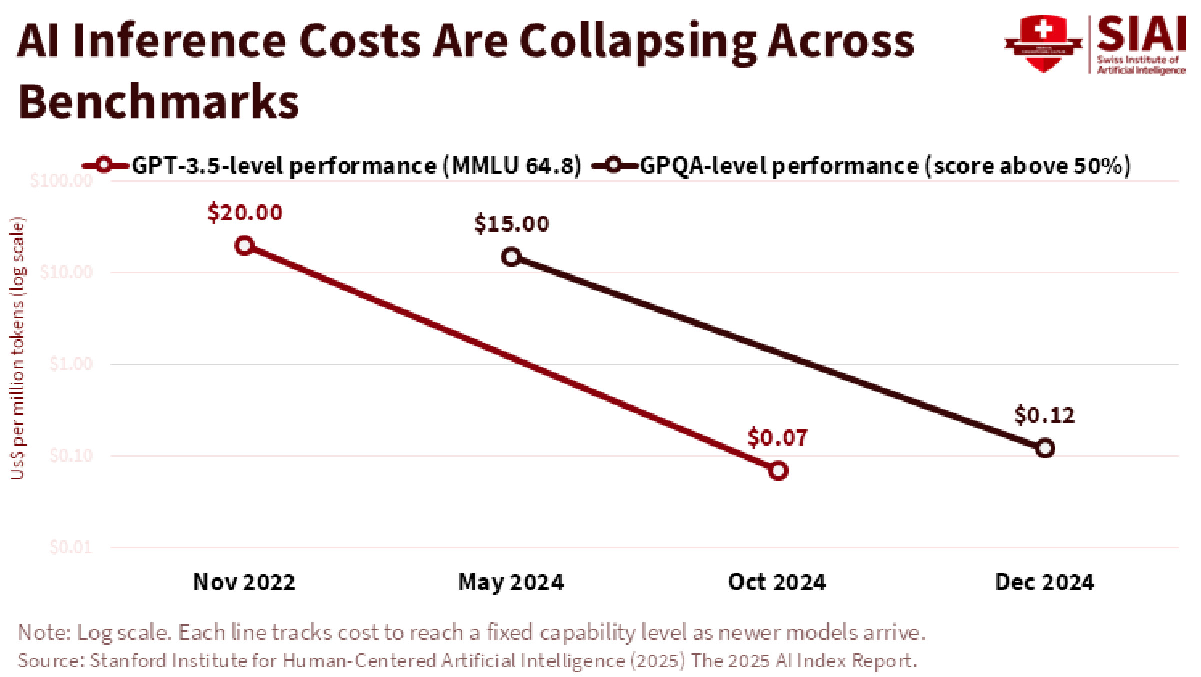


Figure 6: Two different benchmarks, two different timelines, the same collapse.

The evaluation function this paper describes has a cognitive dimension that a purely financial ledger will miss. A 2025 study of 666 participants found a statistically significant negative relationship between frequent AI tool usage and measured critical thinking performance, mediated by the degree to which participants reported offloading cognitive tasks to the tool rather than performing them independently, with the effect most pronounced among younger participants and partially buffered by higher educational attainment.[21] A separate study of 319 knowledge workers led by researchers at Microsoft and Carnegie Mellon found that participants who expressed greater confidence in generative AI reported engaging in less critical thinking when using it for tasks such as developing new ideas or reaching a decision, precisely the tasks where independent judgment matters most.[22] Research on automation bias in AI-assisted decision-making, cited in a recent review of AI overreliance, finds that this suppression of independent evaluation persists across a range of decision scenarios even when the AI system offers no verifiable advantage in accuracy, largely because users struggle to calibrate how much trust a given recommendation deserves.[23] None of this evidence should be read as a claim that cognitive offloading is inherently harmful. Delegating a well-bounded task to a reliable tool is what makes a tool useful and the distinction between ordinary assistance and corrosive over-reliance depends on whether the underlying capability is exercised often enough elsewhere to remain intact and whether the task being delegated is one where an occasional undetected error is tolerable. What the evidence does establish is that this capability erosion is a real cost, not a hypothetical one and that a firm's accounting of AI's value has to include it even though it is harder to price than a token invoice.

The organizations building the governance apparatus to perform this accounting are discovering that it does not sit comfortably inside any single existing department. The OECD's 2025 survey of more than six thousand mid-level managers across France, Germany, Italy, Japan, Spain and the United States showed that managers using algorithmic management tools generally reported improved decision quality but sixty-four percent simultaneously expressed at least one trustworthiness concern, most commonly unclear accountability for algorithmic decisions and an inability to follow the tool's underlying logic.[24] If the people responsible for evaluating AI-assisted work are themselves unsure how much to trust the systems producing the evaluation data, then AI discipline cannot be delegated wholesale to any single function. It requires a joint capability spanning finance, which understands unit costs; operations, which understands workflow design; and domain specialists, who alone can judge whether an output is actually correct. Framing this apparatus as employee surveillance, as some early corporate deployments have done, describes only part of what is being measured. The object being measured is not principally the worker. It is the combined output of a human-machine production system whose behavior neither half can fully explain on its own.

# 4 Workforce Strategy: Distributing AI Dependence Across The Firm

The first two capabilities concern individual roles and individual tasks. The third concerns the aggregate pattern that emerges when many such roles and tasks are combined across a team and then across an organization. Two firms with comparable technology budgets and comparable workforce composition can arrive at very different levels of AI dependence and that difference is not simply a matter of how enthusiastically either firm's employees have taken to the tools. It is substantially a product of governance choices that are made or left unmade, at the organizational level.

Cross-country comparison of otherwise similar economies gives the clearest sign that dependence is a governance variable rather than a technology-access variable. The OECD's 2025 employer survey found that ninety percent of surveyed United States firms had adopted at least one algorithmic management tool to instruct, monitor or evaluate workers, against an average of seventy-nine percent across the four European countries surveyed and just forty percent in Japan.[25] These are advanced economies with broadly comparable access to the same commercial AI products. The gap between them is not explained by which tools exist; it reflects differences in regulatory environment, labor-market institutions and firm-level willingness to build the internal governance needed to deploy such tools responsibly. A separate study applying the Technology-Organization-Environment framework to survey data collected in 2025 found that the depth of AI adoption inside a firm, as distinct from the binary question of whether a firm has adopted AI at all, varies systematically with firm size and available resources, with the relationship between adoption depth and business-model innovation stronger among large firms than among small and medium enterprises.[26] Depth, in other words, is not something that follows automatically from access. It has to be built.

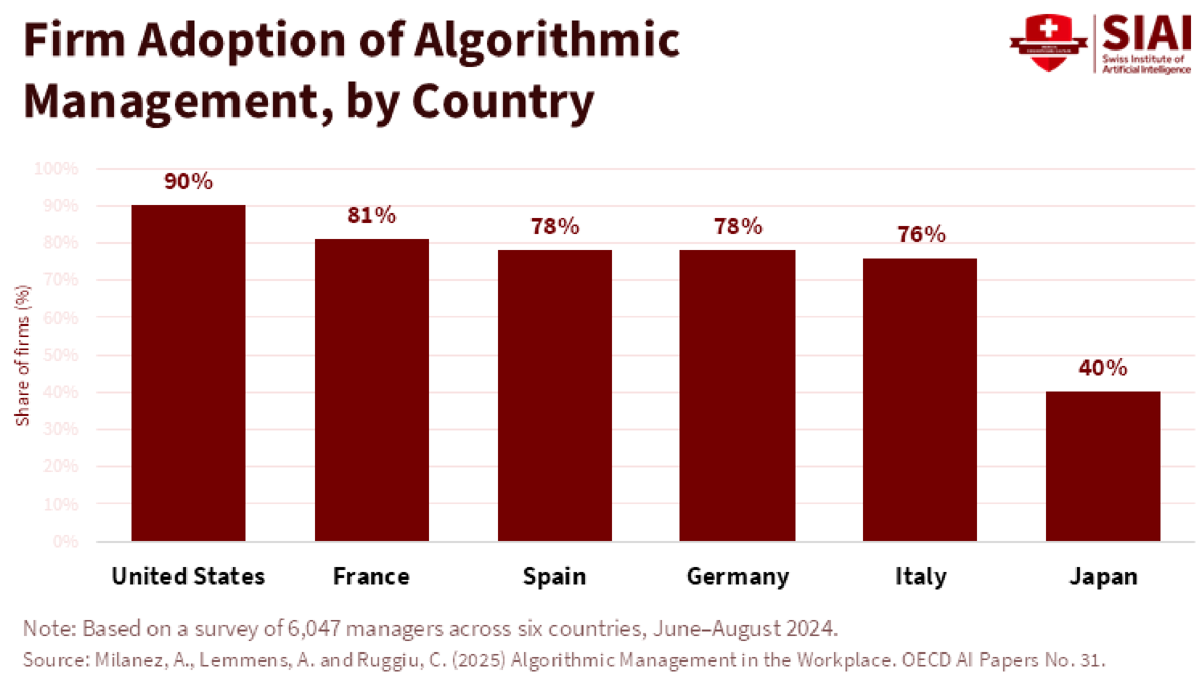


Figure 7: Same technology, same access, more than twice the adoption gap between the U.S. and Japan.

The pace at which this depth is accumulating helps explain why so many firms are managing it poorly. Compiled Eurostat and OECD statistics show enterprise AI use in the European Union rising from 7.7 percent of firms in 2021 to 20.0 percent in 2025, while firm-level AI use across OECD countries with comparable data rose from 8.7 percent in 2023 to 20.2 percent in 2025, an acceleration that has left little time for the governance structures needed to manage it to mature alongside it.[27] A 2026 survey of chief financial officers conducted by the Federal Reserve Bank of Richmond reported that AI-attributed productivity growth, while positive across most sectors, was concentrated well above two percent in high-skill services and finance and considerably more modest elsewhere.[28] That sectoral concentration is itself an argument for differentiated dependence: a marketing function, a legal department, a software engineering group and a compliance office face different jagged frontiers, different error tolerances and different consequences when a model fails silently and there is no reason to expect that the AI operating model suited to one will suit the others.

Evidence from the labor market at large reinforces the case that uniform policy applied across heterogeneous work produces uneven results. One study linked survey-reported ChatGPT use to Danish administrative records across eleven exposed occupations and found essentially no effect on earnings or hours worked through 2024, despite substantial reported usage.[29] A separate analysis of United States data through late 2025 found that 35.9 percent of workers used generative AI regularly by December of that year, alongside small positive wage effects and no statistically significant decline in job openings or employment within exposed occupations.[30] Bharat Chandar's analysis of Current Population Survey data similarly found no aggregate employment decline among high-exposure occupations but with substantial underlying heterogeneity: employment grew in high-education, high-exposure roles such as software development, while it declined in lower-education roles such as customer service.[31] Read together, these findings describe a labor market in which the average effect of AI exposure is close to neutral precisely because sharply divergent effects across occupations are canceling each other out in the aggregate statistics. A workforce strategy built around a single average expectation of what AI will do to a given team will be wrong, in either direction, for most of the specific teams it is applied to.

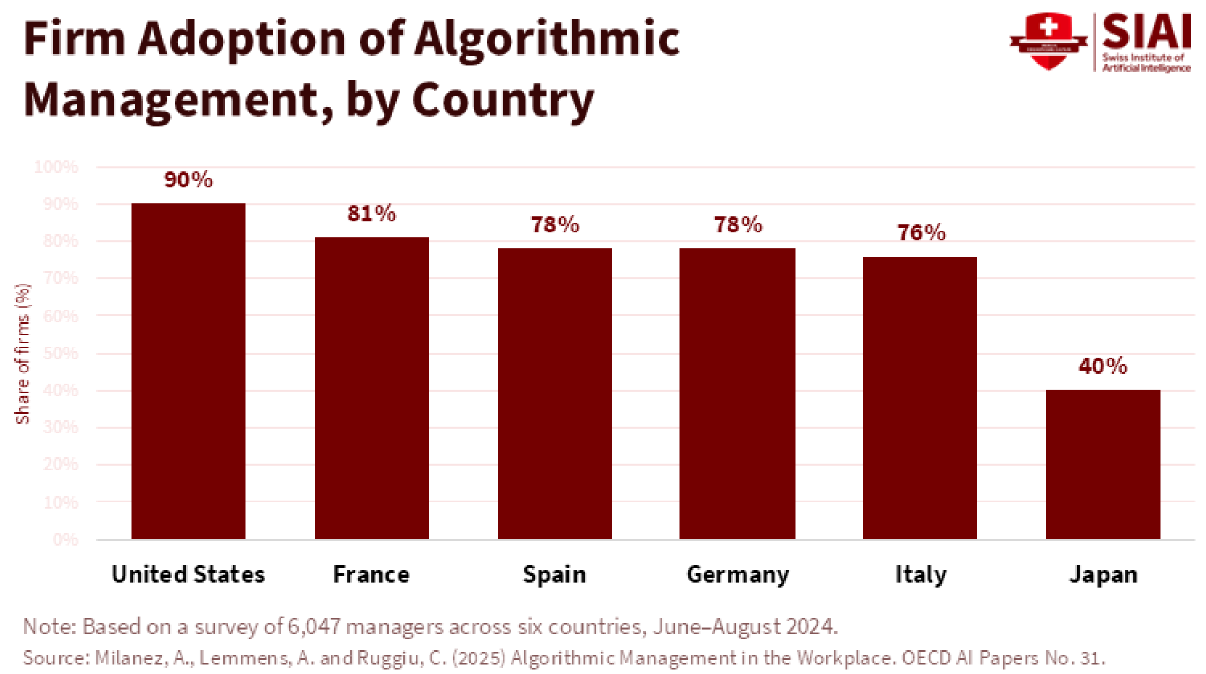


Figure 8: AI is one force among several reshaping employment; the net gain looks reassuring only if you don't look at the 92 million jobs it nets against.

This divergence between formal organizational adoption and actual worker-level usage is itself a governance failure worth naming directly. One study found that, as of February 2024, only 5.4 percent of firms had formally adopted generative AI, a figure that lagged far behind the rate at which individual workers were already using the technology on their own initiative.[32] That gap is the origin of the shadow AI phenomenon documented in MIT's 2025 research, in which employees in the overwhelming majority of firms kept using personal AI accounts regardless of official policy.[33] A firm cannot design an appropriate level of AI dependence for a function it does not know is already AI-dependent in practice. Closing this visibility gap, not simply issuing more licenses or writing a usage policy, is the first practical step toward the kind of workforce architecture this paper is describing, because a policy written without knowledge of actual usage patterns is a policy addressed to an organization that no longer exists.

The strongest counterargument to treating AI dependence cautiously comes from an analysis of how the technology might restructure the labor market for the better rather than the worse. That analysis holds that earlier waves of information technology flattened economic hierarchies in theory but concentrated decision-making authority among elite experts in practice and that AI's distinctive opportunity is to extend some of that higher-stakes decision-making capacity, in fields such as medical care, legal document production and software engineering, to a much larger set of workers who possess the foundational training to use it well.[34] The argument carries real force: if AI genuinely allows a wider set of workers to perform work that previously required scarce specialized expertise, then higher AI dependence in those functions would represent a labor-market improvement rather than a risk to be minimized. The evidence assembled in this paper does not refute that possibility but it does specify its precondition. Autor's own framing makes the benefit contingent on workers possessing the complementary foundational knowledge needed to use AI's output responsibly and on deliberate investment in the training that builds it. That is the organizational capability that separates MIT's five percent of firms extracting real value from the ninety-five percent that are not and it is the same judgment role examined in the first section of this paper. Dependence built on that foundation looks like the productivity gains Brynjolfsson, Noy, Zhang and Peng documented. Dependence built without it looks like the METR slowdown or like the ninety-five percent of stalled pilots that still get counted as adoption.

Translating this into practical responsibility requires naming who should act. Firms bear the primary responsibility for building the joint capability described in the previous section, spanning finance, operations and domain expertise, rather than routing AI governance through human resources alone and for closing the visibility gap between official and shadow AI use before writing dependence policy. Professional bodies in fields such as law, accounting and engineering have a role in establishing occupation-specific standards for what independent verification of AI output should look like, since the jagged frontier differs by domain in ways a generic corporate policy cannot capture. Government intervention is justified where the evidence points to genuine coordination failures rather than ordinary firm-level management choices: the absence of shared standards for disclosing model risk and limitations, measurable underinvestment in retraining for the lower-education, high-exposure occupations where Chandar's data shows employment actually declining and a persistent lack of real-time labor-market data that leaves policymakers and often firms themselves, unable to see how AI dependence is actually distributed across the economy until well after the fact. Assigning firms' internal workforce-architecture decisions to regulators would misplace the responsibility; assigning the training and disclosure infrastructure that surrounds those decisions to firms alone would ignore a genuine collective-action problem that no single employer has the incentive to solve on its own.

There is a further reason this distribution problem cannot be solved once and left alone. The frontier of what AI systems can reliably do is itself moving, which means a dependence architecture calibrated correctly in 2025 may already be miscalibrated by 2027 as model capability, cost and failure modes shift under it. A function that was kept deliberately low-dependence because the available models handled its core judgment tasks poorly may need to be revisited as capability improves, just as a function that was pushed toward high dependence on the strength of early results may need to be pulled back if a provider changes pricing, a model update alters behavior in ways that were not anticipated or accumulated errors surface only after they have compounded across many decisions. Firms that treat their AI operating model as a decision made once, rather than a standing governance function reviewed on a fixed cadence, will find that the architecture they built quietly drifts out of alignment with the technology it was designed around. Workforce strategy for AI differs from earlier waves of technology adoption: the tool itself is non-stationary in a way that a spreadsheet application or a customer relationship management system was not and a governance structure suited to stable tools will not automatically suit one that keeps changing underneath it.

# 5 Conclusion - AI Haves And AI Have-Nots

The boundary between AI haves and have-nots has been described, until recently, in terms of possession: whether a worker or a firm had a Copilot license, an enterprise ChatGPT account or comparable access to a frontier model. That boundary is dissolving quickly and the evidence assembled here suggests it was never the boundary that mattered most. Access has become close to universal within a few years, yet ninety-five percent of the firms that acquired it are seeing no measurable financial return, while employees inside those same firms keep reaching for personal AI accounts that their employer has not sanctioned and cannot see. Possessing the tool and knowing how to organize around it turned out to be almost entirely separate achievements.

The capability that actually distinguishes firms now sits at three levels examined across this paper. The first is managerial: whether an organization has built the judgment, close to the point of production, to know where a jagged frontier of AI competence sits for its own specific tasks and to catch the confident, well-argued error that a challenged model will produce rather than concede. The second is economic and cognitive: whether the organization measures AI's value as a genuine accounting exercise, weighing token cost, review burden and the slow erosion of independent capability against demonstrated output gains, rather than trusting the intuitive sense of speed that experienced professionals have repeatedly been shown to get wrong. The third is architectural: whether AI dependence is deliberately distributed according to the risk and reward of each specific function, rather than adopted uniformly and left to accumulate wherever individual employees happen to push it.

Firms that have built all three capabilities are able to answer a specific question, function by function: how much of this team's output currently depends on AI, at what verified cost, with what fallback if the system fails or a provider changes terms and who is accountable for the answer being right. Firms that cannot answer that question are not meaningfully AI-capable, regardless of how many licenses they have purchased or how enthusiastically their employees have adopted the tools on their own initiative. The next corporate divide will separate organizations that consume AI from organizations that have learned to govern it and the evidence available through 2026 already shows which side of that divide most firms currently occupy.

# References


[1] World Economic Forum (2025) *The Future of Jobs Report 2025*. Geneva: World Economic Forum.

[2] OECD (2023) *OECD Employment Outlook 2023: Artificial Intelligence and the Labor Market*. Paris: OECD Publishing.

[3, 17, 33] Challapally, A., Pease, C., Raskar, R. and Chari, P. (2025) *The GenAI Divide: State of AI in Business 2025*. Cambridge, MA: MIT Project NANDA.

[4] Microsoft (2025) *Work Trend Index 2025*; Mercer (2026) *Global Talent Trends*, as reported in 'Managing Agents Starts Before the Job Title Exists'.

[5] Srinivasan, S. and Wei, V. (2026) 'The Fastest-Growing Job Title of 2026: What an "AI Agent Manager" Actually Does', *The Interview Guys*.

[6] Indeed Hiring Lab (2026) data on AI-referencing job postings, as reported in 'AI Is Coming for Your Job Title'.

[7] Carter, S. (2026) 'The 20 New Agentic AI Jobs Box, McKinsey, And LinkedIn All See Coming', *Forbes*. Reporting LinkedIn's January 2026 workforce analysis.

[8] Sweeney, E. (2025) 'AI is fostering new job titles within HR and people management', *Business Insider*. Reporting Sapient Insights Group's 2025 survey.

[9] Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K.C., Rajendran, S., Krayer, L., Candelon, F. and Lakhani, K.R. (2023) *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality*. Boston: Harvard Business School Working Paper.

[10] Kellogg, K.C., Lifshitz, H., Randazzo, S., Mollick, E., Dell'Acqua, F., McFowland III, E., Candelon, F. and Lakhani, K.R. (2024) *Don't Expect Juniors to Teach Senior Professionals to Use Generative AI: Emerging Technology Risks and Novice AI Risk Mitigation Tactics*. Boston: Harvard Business School Working Paper 24-074.

[11] Randazzo, S., Joshi, A., Kellogg, K.C., Lifshitz, H., Dell'Acqua, F. and Lakhani, K.R. (2026) *GenAI as a Power Persuader: How Professionals Get Persuasion Bombed When They Attempt to Validate LLMs*. Boston: Harvard Business School Working Paper 26-021.

[12] Metaintro (2026) 'AI Manager Is the Hottest New Job Title: How to Break In'.

[13] Brynjolfsson, E., Li, D. and Raymond, L.R. (2023) *Generative AI at Work*. Cambridge, MA: NBER Working Paper 31161. Published (2025) in *Quarterly Journal of Economics*, 140(2), pp. 889–942.

[14] Noy, S. and Zhang, W. (2023) randomised trial on generative AI and professional writing productivity, as summarised in NBER Reporter (2024) 'The Economics of Generative AI'.

[15] Peng, S., Kalliamvakou, E., Cihon, P. and Demirer, M. (2023) field study of GitHub Copilot and software development time, as summarised in International Center for Law and Economics (2026) *AI, Productivity, and Labor Markets: A Review of the Empirical Evidence*.

[16] Becker, J., Rush, N., Barnes, E. and Rein, D. (2025) *Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity*. METR.

[18] Menlo Ventures (2026) enterprise generative AI spending estimates, as compiled in industry FinOps-for-AI cost analyses.

[19] Stanford Institute for Human-Centered Artificial Intelligence (2025) *The 2025 AI Index Report*. Stanford, CA: Stanford University.

[20] FinOps Foundation (2026) *State of FinOps 2026*, as reported in Virtasant 'FinOps for AI: Creating Token-Level Visibility for Practitioners'.

[21] Gerlich, M. (2025) 'AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking', *Societies*, 15(1), article 6.

[22] Microsoft and Carnegie Mellon University (2025) study of 319 knowledge workers, *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, as summarised in American Psychological Association Monitor (2026) 'How AI is reshaping human skills and thinking'.

[23] Klingbeil, A. et al. (2024) findings on automation bias in AI-assisted decision-making, as reported in ScienceDirect (2026) 'AI-overdependence and human cognitive decline: Hazards, evidence, and mitigation strategies'.

[24, 25] Milanez, A., Lemmens, A. and Ruggiu, C. (2025) *Algorithmic Management in the Workplace: New Evidence from an OECD Employer Survey*. Paris: OECD Artificial Intelligence Papers No. 31.

[26] Study applying the Technology-Organization-Environment framework to AI adoption depth across firm sizes (2025) *Journal of Asian Business Studies*, 19(4).

[27] Compiled Eurostat and OECD firm-level AI adoption statistics, as reported in 'The Agentic Economy: Humans, AI Agents, Robots, and the Measurable Transition toward Distributed Economic Action' (2026).

[28] Federal Reserve Bank of Richmond (2026) CFO survey on AI-attributed productivity growth by sector, 'Artificial Intelligence, Productivity, and the Workforce'.

[29] Humlum, A. and Vestergaard, E. (2025) study linking ChatGPT use to Danish administrative earnings and hours records, as summarised in International Center for Law and Economics (2026) *AI, Productivity, and Labor Markets: A Review of the Empirical Evidence*.

[30] Hartley, J.S., Jolevski, F., Melo, V. and Moore, B. (2026) analysis of United States generative AI usage and employment outcomes through December 2025, as summarised in International Center for Law and Economics (2026) *AI, Productivity, and Labor Markets: A Review of the Empirical Evidence*.

[31] Chandar, B. (2025) Current Population Survey analysis of AI exposure and occupational employment outcomes, as summarised in International Center for Law and Economics (2026) *AI, Productivity, and Labor Markets: A Review of the Empirical Evidence*.

[32] Bick, A., Blandin, A. and Deming, D.J. (2024) *The Rapid Adoption of Generative AI*. Cambridge, MA: NBER Working Paper 32966.

[34] Autor, D.H. (2024) *Applying AI to Rebuild Middle Class Jobs*. Cambridge, MA: NBER Working Paper 32140.