Scholedge International Journal of *Management & Development*
ISSN 2394-3378 Vol. 11, Issue 02 (2024)
DOI: https://dx.doi.org/10.19085/sijmd110201

AI Adoption and the Trust Problem: What Managers Need to Know About Employee Skepticism

Oluwafemi Bello
University of Maiduguri, Maiduguri, Nigeria

ABSTRACT

Artificial intelligence tools have moved from the margins of organizational life to the center of everyday work. Yet the speed of that move has outpaced the willingness of many employees to genuinely trust and use these tools for their official responsibilities. This paper examines the nature of that trust gap, drawing on organizational trust theory, human-automation interaction research, and publicly documented real-world cases. The analysis shows that employee skepticism about AI does not stem from ignorance or stubbornness. It comes from legitimate worries about job security, surveillance, lack of transparency, and a history of AI systems that have failed to deliver what was promised. The paper then presents seven concrete advisories for managers seeking to address these concerns, along with specific guidance on how to implement each one in day-to-day organizational practice. The goal is not to persuade employees to simply accept AI, but to create the conditions under which trust can develop honestly and on solid ground.

Keywords: Artificial intelligence, workplace trust, employee skepticism, AI adoption, organizational management, psychological safety

INTRODUCTION

Few technologies have entered the workplace as quickly, and as noisily, as artificial intelligence. Within a couple of years of generative AI tools becoming widely accessible, organizations across sectors began asking their employees to use these systems for drafting documents, analyzing data, summarizing reports, responding to queries, and dozens of other tasks that once required pure human effort. The business case for this shift is not hard to understand. A McKinsey & Company (2023) survey found that one-third of organizations were already regularly using generative AI in at least one business function, and more than two-thirds expected to increase AI investment over the next three years.

But adoption figures and investment projections tell only part of the story. The other part involves what employees actually think about the AI tools being placed in front of them, and whether they trust those tools enough to rely on them for work that carries real professional consequences. That question matters enormously, because technology that sits unused or is used reluctantly produces very little of the value organizations expect from it.

The trust problem around AI in workplaces is not a new observation, but it has grown more pressing as AI tools have moved from specialized back-office applications to general-purpose assistants that employees are expected to use daily. Workers are being asked to delegate parts of their cognition to systems they do not fully understand, whose reasoning is rarely explained to them, and whose presence they sometimes associate with surveillance, deskilling, or eventual job replacement. These are not irrational fears, and treating them as such is one of the more common mistakes managers make when rolling out AI tools.

This paper takes the position that managing the AI trust deficit is one of the most important leadership challenges facing organizations today. It begins by grounding the discussion in what organizational research tells us about trust, then examines the specific mechanisms through which AI erodes or prevents trust from forming among employees. It draws on real-world cases, including Amazon and IBM, to show what the trust deficit looks like in practice. It concludes with a set of actionable advisories for managers, each accompanied by practical implementation guidance that goes beyond the usual platitudes about communication and training.

UNDERSTANDING ORGANIZATIONAL TRUST: WHAT THE RESEARCH TELLS US

Before it is possible to understand why employees distrust AI, it is worth understanding what trust in an organizational context actually means. The most widely cited framework in this area comes from Mayer, Davis, and Schoorman (1995), who defined trust as the willingness of a party to be vulnerable to the actions of another party, based on the expectation that the other party will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party. In their model, trust depends on three perceptions of the trustee: ability (competence to do what is needed), benevolence (genuine care for the trustor's interests), and integrity (adherence to an acceptable set of principles).

When employees evaluate a new AI tool, they apply something very close to this framework, often unconsciously. They ask: Is this system actually competent at the tasks it is being asked to perform? Does it have my interests in mind, or only the organization's? And does it operate according to values I can accept? In many workplace AI deployments, at least one of these three dimensions is missing or unclear, and that gap is enough to produce skepticism even among employees who are otherwise open to new technology.

Lee and See (2004), in their influential review of trust in automation, pointed out that trust in technology follows patterns similar to trust in people. People assess automated systems over time, developing calibrated trust based on whether the system performs reliably across different situations. Crucially, they noted that the goal is not maximum trust, but appropriate trust: enough trust to rely on the system when it is genuinely helpful, and enough skepticism to catch it when it is not. Over-trust is just as problematic as under-trust, because an employee who blindly accepts AI outputs without critical reflection can propagate errors as easily as one who ignores the AI entirely.

There is also an important dimension of organizational climate to consider. Edmondson (1999) established that psychological safety, the shared belief within a team that it is safe to take interpersonal risks, is a prerequisite for learning behavior in organizations. When employees do not feel psychologically safe, they hide mistakes, avoid experimentation, and disengage from processes they find threatening. Workplaces with low psychological safety are particularly hostile environments for AI adoption, because using a new and imperfect tool necessarily involves making mistakes, asking basic questions, and admitting uncertainty. If the organizational culture punishes those behaviors, employees will either avoid the AI or use it in ways that look compliant but are actually performative.

Taken together, this body of research suggests that trust in AI at work is not primarily a technology problem. It is an organizational and relational problem. Employees need to perceive the AI as competent, believe it operates in their interest, see its values as compatible with their own, and exist within a culture that supports genuine experimentation. When any of those conditions is absent, skepticism fills the gap.

WHY EMPLOYEES ARE SKEPTICAL: THE SOURCES OF THE TRUST DEFICIT

1. Fear of Job Displacement

The most commonly cited source of employee skepticism about AI is the fear of being replaced. This is not a paranoid fantasy. McKinsey's research on generative AI and the future of work found that automation could affect roughly 30 percent of work hours in the United States economy by 2030, and that workers in lower-wage brackets were up to 14 times more likely to need to change occupations than those in higher-wage positions (McKinsey & Company, 2023). A separate McKinsey survey on workforce concerns found that 35 percent of employees cited workforce displacement as a concern around generative AI, alongside 51 percent who worried about cybersecurity and 50 percent who were concerned about inaccuracy (McKinsey & Company, 2025).

Those numbers tell a more layered story: displacement anxiety is not the only concern, and it is not even the most common one. But it is the concern that most directly affects how employees engage with the AI tools placed in front of them. An employee who believes that doing their job well with AI assistance will ultimately result in their role being eliminated has every rational reason to use that AI reluctantly, if at all. The tool that is supposed to make them more productive can feel, from their vantage point, like the thing that is actively building the case for their redundancy.

Managers who dismiss these fears as unfounded rarely succeed in dispelling them. The concern has enough documented precedent that it cannot be waved away with reassurances that AI will only create new jobs. A more effective response is to be specific and honest about what the organization's AI adoption will and will not mean for individual roles, which the advisories in Section 5 address directly.

2. Surveillance and Autonomy Concerns

A second major source of skepticism comes from the way AI is sometimes deployed not as a tool to help employees but as a mechanism to monitor them. Performance tracking, productivity scoring, email analysis, and activity logging systems have become common features of AI-enhanced workplaces, and employees often experience these tools very differently from how management intends them. Even when the stated purpose is operational efficiency or quality assurance, employees frequently perceive such systems as instruments of control rather than support.

Trust is inherently reciprocal, and this is why surveillance cuts so deep. When employees feel watched in ways they find disproportionate or opaque, their trust in the employing organization drops, and that drop extends to the AI systems associated with the surveillance. Organizations that introduce monitoring AI without employee input, without transparency about what is being tracked and why, and without clear limits on how data will be used tend to produce exactly the kind of adversarial relationship that makes subsequent AI adoption difficult.

The dynamic also has a practical downside for organizations. Research reviewed by the Center for Democracy and Technology found that electronic monitoring does not reliably improve employee performance and may actually lead to counterproductive work behaviors (as cited in The Register, 2025). Employees who feel surveilled invest energy in appearing productive rather than being productive, and they learn to game the metrics rather than improve their actual work.

3. Lack of Transparency and Explainability

Many AI systems, particularly those built on machine learning, function in ways that are genuinely difficult to explain. When a system makes a recommendation, flags a piece of work, or scores a performance review, it often cannot offer a clear account of why it reached that conclusion. This opacity is a structural feature of many current AI architectures, not a design choice that organizations can simply override.

For employees, working with a system that cannot explain itself raises obvious problems. If an AI tool recommends a particular course of action in a customer interaction or a legal document, and the employee cannot understand the basis for that recommendation, they face an uncomfortable choice: accept the output uncritically, which feels like an abdication of professional judgment, or ignore it, which defeats the purpose of using the tool. Neither option builds trust, and both options leave the employee in a position of feeling responsible for outcomes they did not fully understand or control.

McKinsey's 2023 survey noted that just 21 percent of organizations had established policies governing employees' use of generative AI technologies in their work (McKinsey & Company, 2023). In the absence of such policies, employees are left to navigate the ambiguity themselves, which tends to produce inconsistent and often anxious patterns of use. Some employees use the AI heavily but feel they cannot say so openly. Others avoid it entirely to protect themselves from being associated with errors they cannot predict.

Davenport and Ronanki (2018) pointed out, in their early examination of enterprise AI adoption, that AI rollouts fail most visibly when organizations underestimate the importance of explaining what the technology does and does not do to the people expected to use it. The gap between the capabilities organizations communicate to employees and the capabilities the tools actually possess is itself a source of distrust, because it creates situations where employees feel they were misled.

4. Past AI Failures and Institutional Memory

Employees are not naive about technology. Many of them have lived through cycles of technological promises that did not materialize, or systems that were introduced with fanfare and then quietly abandoned after causing disruption. This institutional memory shapes how they respond to new AI tools, often with a skepticism that managers may find baffling but that makes perfect sense in context.

There is also the broader public record of AI failures to draw on. When employees read about AI systems that produced discriminatory outcomes in hiring or lending, or that failed in high-stakes medical contexts, that information quietly shapes the expectations they bring to the AI tools their own organization is asking them to use. The failures do not have to be in their own industry to shape their expectations.

REAL-WORLD CASES: WHEN THE TRUST DEFICIT APPEARS

1. Amazon and the Automation of Discipline

One of the most extensively documented cases of AI creating a trust breakdown between an organization and its workforce is Amazon's use of algorithmic management in its warehouse operations. Amazon developed an internal system known as the Associate Development and Performance Tracker (ADAPT), which monitored worker productivity in real time, tracked Time Off Task with precision measured in minutes, and could automatically generate warnings and termination notices without requiring input from a human supervisor (as reported in various accounts, including the Berkeley School of Information, 2022, and UNI Global Union, 2023).

The results, from an employee trust perspective, were severe. A survey commissioned by UNI Global Union in 2023, covering 2,000 self-identified Amazon workers across eight countries, found that Amazon's monitoring system made workers feel stressed, pressured, anxious, and untrusted. More than half reported negative impacts on their physical health, and 57 percent reported negative impacts on their mental health. Drivers were particularly affected, with 65.7 percent reporting a negative physical health impact directly tied to productivity monitoring.

What the Amazon case illustrates is not simply that monitoring is bad for morale. It illustrates a more specific mechanism: when employees experience AI as something that is done to them rather than for them,

the result is not compliance but resentment, withdrawal, and a generalized erosion of trust in the employing organization. Workers at Amazon did not respond by working harder. Many worked in ways that gamed the metrics, protected themselves from the system rather than engaging honestly with it, and increasingly sought collective responses through unionization.

Amazon's case also shows what happens when a quantified performance metric becomes the primary lens through which management views employee behavior. The system created no space for individual circumstances, no room for nuance, and no human judgment at the point of discipline. Employees who experienced this described feeling like the company viewed them as robots, a phrase that recurred in multiple independent accounts. Trust does not survive that kind of relationship for long.

2. IBM Watson Health and the Cost of Overpromising

A different kind of trust problem emerges in the case of IBM Watson Health. IBM invested heavily in positioning Watson as a system that could transform clinical decision-making, particularly in oncology. The marketing around Watson's capabilities was extraordinarily bold: IBM's CEO described it as the company's healthcare moonshot, and the public messaging implied that Watson would soon be capable of helping cure cancer by analyzing vast bodies of medical literature and patient data.

In practice, Watson for Oncology struggled to deliver reliable, culturally applicable recommendations. Reports from clinical partners revealed that Watson's suggestions were often based on U.S.-centric treatment guidelines that were difficult to apply in other healthcare systems, and that the training data included hypothetical cases rather than real patient outcomes. By 2018, skepticism among oncologists had grown substantially, with some reporting that Watson's recommendations conflicted with established clinical practice and that the system's interface disrupted rather than supported their workflow. As noted in IEEE Spectrum (2019), some physicians resisted the tool precisely because they trusted their own clinical judgment more than a black box recommendation from a machine they did not understand.

IBM sold Watson Health's assets in 2022 for approximately one billion dollars, a fraction of the billions it had invested in building and acquiring the division. The failure was not purely technical. A significant component was trust: trust from clinical staff who felt the system was not built with genuine understanding of their work, trust from institutional partners who felt the marketing had misrepresented what the tool could do, and trust from the internal IBM teams who had staked their professional credibility on a product that could not live up to the vision they had been asked to sell.

The lesson for organizations is that trust erosion caused by overpromising is particularly hard to repair, because it combines disappointment with the feeling of having been misled. Employees and professional users who believe they were sold a false picture of what AI can do do not simply revise their expectations. They become broadly skeptical of all subsequent AI claims, which makes future adoption more difficult and more expensive in terms of both time and organizational goodwill.

3. Bossware and the Remote Work Backlash

The COVID-19 pandemic accelerated a third pattern that has continued to generate trust problems. As organizations scrambled to manage remote workforces, many turned to AI-assisted employee monitoring tools, sometimes called bossware, to track activity levels, log keystrokes, capture screenshots, and analyze communication patterns. A PwC survey noted that by January 2021, one in five companies was using some form of surveillance software to monitor remote employees, in some cases without the employees' knowledge.

The response from employees was largely negative, and not only for the obvious reasons of privacy. The deeper problem was what the monitoring communicated about organizational values. Employees who

had successfully managed their own workloads for months, sometimes producing better results than they had in the office, found themselves subjected to surveillance that felt designed for distrust rather than support. The implicit message, regardless of how management framed it, was: we do not believe you are actually working.

The bossware problem is directly connected to broader AI trust because bossware experiences have shaped the expectations many employees now bring to every AI tool their organization deploys. Employees who have experienced opaque monitoring do not start from a position of good faith when a new AI system is introduced. They ask, rightly, what this new tool is measuring, who sees the data, and how it will be used against them. Organizations that have not addressed those questions clearly face a much harder road to genuine AI adoption.

MANAGERIAL ADVISORIES: BUILDING AND SUSTAINING EMPLOYEE TRUST IN AI

The following advisories are intended to give managers concrete direction for addressing the trust deficit around AI. Each is grounded in the evidence reviewed in this paper, and each is accompanied by specific guidance on implementation. The advisories are not a checklist to be completed once. They describe an ongoing approach that needs to be woven into how organizations manage their relationship with employees over time.

Advisory 1: Be Honest About What the AI Does and Cannot Do

The single most common source of AI trust failure is the gap between what employees are told AI will do and what it actually delivers. Managers who inherit vendor claims and pass them along uncritically, or who frame AI tools in aspirational terms to generate enthusiasm, often create the conditions for exactly the disillusionment they were trying to prevent.

The more durable approach is to be genuinely straightforward from the beginning. Before deploying any AI tool, managers should communicate clearly: what tasks the tool is designed to assist with, what its known limitations are, what kinds of errors it tends to make, and what the organization expects employees to do when the tool is wrong. This kind of communication treats employees as professionals capable of handling nuanced information, which is itself a trust signal. It also sets realistic expectations that the tool is more likely to meet, which builds credibility over time.

Prepare a one-page plain-language overview of each AI tool before deployment, covering its purpose, its limitations, and the procedures for flagging problems. Share this in a team meeting rather than through email, so employees have the opportunity to ask questions in real time. Revisit the overview after the first month of use to update it based on what has actually happened.

Advisory 2: Involve Employees Before Deployment, Not After

Most organizations introduce AI tools as a fait accompli, telling employees that starting from a certain date they will be using a new system. This approach skips the step that most determines whether employees will engage with the tool genuinely or performatively: the step at which they have some say in how it will be used.

Involving employees in the AI deployment process does not mean giving them a veto over organizational decisions. It means treating them as people with relevant expertise about their own work, which they are. Employees who process insurance claims all day know things about that task that external consultants and technology vendors do not, and the AI tool that accounts for their knowledge from the start will be better for it, and more trusted. Employees who feel they had a voice in the decision are also substantially more likely to engage with the outcome honestly, including reporting problems when the tool does not work as expected.

At least six weeks before deployment, convene a small working group of frontline employees who will actually use the tool. Give them a sandbox version to test and ask them to document situations where it performs well and situations where it does not. Use their observations to shape how the tool is configured, what guardrails are put in place, and how training is designed. Acknowledge their contributions explicitly when deployment begins.

Advisory 3: Establish Clear Policies on Data, Privacy, and Performance

One of the key reasons employees distrust AI tools is uncertainty about what happens to the data those tools generate. When an employee uses an AI writing assistant or interacts with a customer service AI, they reasonably wonder: is this interaction being logged? Who has access to the logs? Can it be used to evaluate my performance? Could it eventually be used to argue for replacing my role?

The absence of clear, written answers to these questions does not make employees less worried. It makes them more worried, because they have to imagine the answers themselves, and people tend to imagine the worst when they are not given facts. Organizations that establish and communicate explicit data governance policies for their AI tools remove a significant source of anxiety, and that removal has a direct positive effect on trust.

McKinsey's 2023 survey found that only 21 percent of organizations using generative AI had established policies governing employee use of those tools. This is a remarkable gap, and it suggests that most organizations are asking employees to make judgment calls about AI use without giving them the framework to make those calls well.

Before any AI tool is deployed, the organization should document the answers to four questions: What data does this tool collect? Who has access to it? For how long is it retained? And will it be used in any way to evaluate individual employee performance? These answers should be communicated in writing, reviewed by HR and legal counsel, and made available in a place employees can easily find. If data use policies change, employees should be notified proactively.

Advisory 4: Reframe AI as a Tool, Not a Judge

A great deal of the anxiety employees feel about AI tools comes from the way those tools are positioned. When AI is described in terms of efficiency, optimization, and performance measurement, employees hear the subtext: the AI is watching whether you are doing your job well enough. When AI is described as a collaborator that handles tedious tasks so employees can focus on the parts of their work that actually require human judgment, the same tools produce a very different emotional response.

This is not simply a matter of communication strategy or rebranding. The way managers genuinely think about AI, and the way they discuss it in meetings, performance reviews, and casual conversations, shapes the organizational culture around it. Managers who treat AI as a source of surveillance data, even implicitly, will produce a culture in which employees treat it as a threat. Managers who treat it as a resource for problem-solving will produce a culture in which employees are more willing to experiment with it honestly.

The distinction Davenport and Ronanki (2018) drew between AI systems designed to augment human work and those designed to automate it entirely is useful here. Augmentation-focused AI, properly explained and positioned, tends to generate more employee acceptance, because it affirms rather than threatens the value of human expertise.

Review all organizational messaging about AI tools before deployment. Remove or revise language that emphasizes performance monitoring, efficiency gains at the expense of headcount, or the eventual reduction of human involvement. Replace it with language that is specific about what human capabilities

the AI is meant to support, and where the organization sees continued human judgment as irreplaceable. Be consistent in this framing across all communication channels.

Advisory 5: Create Real Channels for Employees to Report Problems

Employees who use AI tools every day will notice failures that management and vendors will not catch in testing. A customer service AI that handles standard queries perfectly but consistently misreads frustrated or emotionally distressed customers is a problem that shows up in the lived experience of frontline staff long before it shows up in performance dashboards. But employees will only report these problems if they believe doing so will result in something useful rather than defensive denial or, worse, a sense that they are being seen as resistant to change.

Edmondson's (1999) work on psychological safety is directly relevant here. Organizations in which employees feel safe to raise concerns about AI are organizations in which AI tools improve over time, because the feedback loop between human experience and system design actually functions. Organizations in which employees feel they will be dismissed or penalized for raising concerns end up with AI tools that quietly continue to perform poorly, because no one with useful information feels it is worth the risk to say so.

Establish a structured feedback mechanism for each AI tool within the first month of deployment. This could be as simple as a monthly team meeting with a standing agenda item for AI performance issues, or a dedicated channel in the organization's communication platform for AI-specific feedback. More importantly, managers must be seen responding to that feedback. When an employee raises a concern and nothing happens, the channel dies. When a concern is raised and addressed within a reasonable timeframe, trust in the reporting process builds.

Advisory 6: Invest in Meaningful Training, Not Just Technical Orientation

Most AI training in organizations focuses on how to use the tool: how to write a good prompt, how to navigate the interface, how to export outputs. This kind of training is necessary but insufficient. It does not address the harder questions that shape how employees actually engage with AI day to day: when should I trust what the AI tells me? When should I override it? What am I responsible for when I use an AI-assisted output in a client-facing context? How do I maintain my own professional judgment rather than outsourcing it?

These are not abstract philosophical questions. They are practical decisions employees face constantly once AI tools are part of their workflow, and employees who have not thought them through in a supported setting tend to handle them inconsistently, sometimes with significant professional or organizational risk. Lee and See (2004) described the goal of appropriate trust calibration as knowing when to rely on automation and when not to. That calibration does not happen automatically. It requires deliberate reflection on the nature of the tool and the nature of the task, and that reflection is best structured through training.

Beyond technical orientation, design a two-hour session for each major AI tool deployment focused specifically on judgment. Use real scenarios drawn from the employees' own work, present them with AI outputs that range from excellent to subtly wrong to clearly mistaken, and facilitate discussion about how employees would evaluate and respond to each. Make the session interactive rather than didactic, and make it clear that there are no wrong answers as long as the reasoning is transparent.

Advisory 7: Treat Trust as an Ongoing Project, Not a Launch Event

The biggest structural mistake organizations make with AI adoption is treating it as a one-time event followed by normal operations. A tool is introduced, training is completed, and employees are expected to

use it. If problems arise, they are handled case by case. But trust in AI tools, like trust in any relationship, is not static. It builds or erodes based on ongoing experience, and it requires ongoing attention from management.

Employees who have a bad experience with an AI tool and never receive any follow-up from management carry that experience forward. Employees who see AI tools updated in response to the concerns they raised feel that their experience matters. Employees who are periodically asked how the tool is working, and who see those conversations lead to real changes, develop a different relationship to AI adoption over time. The compound effect of consistent managerial attention to the trust dimension of AI use is substantially larger than any individual intervention.

Mayer et al. (1995) noted that trust develops through repeated interactions that confirm or disconfirm expectations. This applies directly to AI in the workplace: every time an AI tool delivers what was promised, trust edges upward. Every time it fails without any organizational response, trust erodes. The manager's job is to shape the context in which those interactions occur, to ensure that failures are acknowledged honestly and addressed where possible, and to communicate consistently that employee experience with AI is taken seriously.

Schedule a quarterly review of each major AI tool's performance from the employee perspective. This is separate from technical performance metrics. It asks: Are employees using the tool? Are they reporting that it helps? What problems have they encountered? Have those problems been addressed? What would make the tool more useful? Feed the answers into an ongoing dialogue with the tool vendor, with HR, and with senior leadership, and communicate the outcomes of that dialogue back to employees. This cycle, maintained over time, is what transforms reluctant compliance into genuine adoption.

CONCLUSION

The trust deficit around AI in the workplace is real, it is widespread, and it is not going away on its own. Employees who are asked to rely on AI tools for official tasks are making a judgment call about whether those tools are competent, whether they serve employee interests, and whether the organizations deploying them can be trusted to use AI ethically. When those judgments land on the skeptical side, the result is not a smooth adoption process with a few bumps. The result is a hidden resistance that costs organizations the value they expected to capture from their AI investment while simultaneously damaging the organizational relationship with employees.

The evidence from organizational trust research, from human-automation interaction studies, and from cases like Amazon and IBM Watson Health all points in the same direction. Trust in AI at work is not primarily a technology problem. It is an organizational problem, and it requires organizational solutions. Managers who treat the trust deficit as a communication challenge to be solved with better rollout messaging are likely to find that the messaging does not stick. What actually works is addressing the conditions that produce distrust in the first place: displacement anxiety, surveillance, opacity, and the absence of genuine employee voice in AI decisions.

The advisories presented in this paper do not guarantee that employees will embrace AI enthusiastically. Trust is not something that can be engineered or mandated. What the advisories offer is a framework for creating the conditions under which trust can develop honestly, through repeated experience with AI tools that perform as described, in organizational environments where employees feel their concerns are heard and their professional judgment is valued. That process takes time, but it is the only process that produces lasting results.

For managers, the practical takeaway is this: the work of AI adoption does not end when the tools are deployed. In many ways, it begins there. The organizations that will get the most out of AI over the next

decade are not necessarily the ones with the most sophisticated tools. They are the ones that take the human side of that equation seriously, and who understand that earning the trust of their employees is at least as important as selecting the right algorithm.

REFERENCES

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96, 108–116.

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. Administrative Science Quarterly, 44(2), 350–383. https://doi.org/10.2307/2666999

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

McKinsey & Company. (2023). The state of AI in 2023: Generative AI's breakout year. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year

McKinsey & Company. (2025). Superagency in the workplace: Empowering people to unlock AI's full potential at work. https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of organizational trust: Past, present, and future. Academy of Management Review, 32(2), 344–354. https://doi.org/10.5465/amr.2007.24348410

UNI Global Union. (2023). New survey: Amazon workers in 8 countries say intrusive monitoring is making them sick and anxious. https://uniglobalunion.org/news/globalsurvey23/