Policy, Not Public Relations: What Actually Drives Public Distrust of Artificial Intelligence

The Economy Research Editorial*

*The Economy Research, 71 Lower Baggot Street, Dublin 2, Co. Dublin, D02 P593, Ireland

Abstract

Public trust in AI keeps sliding even as more people use it every day and that gap gets overlooked. A 2026 Brookings analysis attributes the slide to bad marketing from AI companies and calls for stricter rules on how AI gets advertised, but that gets the story backward: firms selling AI have little reason to make their own product look worse and if anything, they're doing the opposite. This paper argues the real driver is more ordinary and a lot less exciting: current AI systems are still wrong often enough to notice and many firms rolled out AI customer service faster than the technology could actually handle, cutting costs before the tool was ready. Trustpilot's review data, Gartner's own reversal on staffing cuts and Pew's polling all point the same way. On the labor side, AI is reshaping work rather than wiping it out, but the reshaping so far rewards people who already know how to supervise AI output and leaves almost no room for beginners to learn on the job. None of this is fixed by regulating what AI companies say about themselves. It is fixed by how firms actually deploy the tool and by how fast new entry-level paths into AI-supervised work open up.

1 Introduction — Why AI's Trust Problem Isn't a Marketing Problem

When a customer contacts an insurer after a house fire and is routed through three automated menus before reaching no one, the injury is not abstract. It shows up as a number: insurers mentioned in AI-related reviews on Trustpilot's platform averaged 2.47 stars between 2021 and 2025, compared with 4.35 stars for reviews that never mentioned AI at all and a single unresolved, unescalated complaint was enough to cost a company an average of 1.6 stars.[1] That gap is not a communications problem in the sense the public relations industry usually means. No slogan closes it, because the thing producing the low rating is not a message. It is the software and the way a firm chose to deploy it.

This distinction matters because the dominant explanation now circulating in policy circles gets the causal order backward. A widely discussed 2026 Brookings analysis argues that young people's distrust of artificial intelligence stems substantially from how AI companies have marketed the technology and that credible policy, not public relations, is what will ultimately earn Generation Z's trust.[2] The argument has real elements: labor protection gaps, unresolved questions about compensation for creative labor and opaque data center siting decisions are all legitimate objects of regulation. But treating marketing and policy failure as the primary explanation for public distrust confuses a downstream symptom with the underlying cause. It also assumes a strategic alignment among AI vendors that does not obviously exist. Firms selling generative AI systems are engaged in an unusually intense competition for market share, revenue and enterprise contracts; there is little commercial incentive for any of them to cultivate a negative public image and considerable incentive to do the opposite. If AI's reputation has nonetheless deteriorated, that deterioration is better explained by what the technology and its deployment actually do to people, day after day, in ordinary transactions, than by any coordinated messaging failure.

The scale of the underlying shift makes this an urgent question rather than a merely academic one. In February 2026, Pew Research Center found that 49 percent of American adults had used an AI chatbot, up sharply from 33 percent two years earlier and that 44 percent had specifically used ChatGPT, more than double the 2023 figure.[3] Adoption, in other words, has moved from early enthusiasm to near-majority mainstream use in roughly three years. Trust has not kept pace. In the same period, half of American adults reported feeling more concerned than excited about AI's growing role in daily life, up from 37 percent in 2021 and only about a quarter expected AI to improve education or the way people do their jobs.[4] This is not a generational phenomenon confined to Gen Z, whatever the framing of any single report might suggest; distrust runs across age cohorts, occupations and political affiliations, even as usage climbs in all of them. What has changed since 2023 is not public relations. It is that tens of millions of people now interact with AI systems directly and repeatedly enough to form their own judgments about reliability, independent of what any company says about its product.

This paper argues that the source of public distrust in AI is neither primarily a marketing failure correctable through better messaging, nor primarily a policy vacuum correctable through advertising standards but a structural mismatch between what current AI systems reliably do and what the tasks people delegate to them

require. That mismatch shows up most visibly in customer-facing deployments, where firms have used AI to cut costs faster than the technology's actual competence allows, producing measurable reputational damage that is now beginning to reverse through rehiring.[5] It shows up in the persistent gap between the ease of cognitive offloading and the retained obligation to exercise judgment, verify outputs and accept responsibility for decisions, an obligation that generative systems, unlike calculators or search engines, make unusually easy to abdicate rather than merely defer.[6] And it shows up in a labor market that is visibly reorganizing around AI but unevenly, in ways that concentrate opportunity among experienced workers while leaving entry-level pathways underdeveloped.[7] None of these dynamics will be resolved by regulating what AI companies are permitted to say about their products. They will be addressed, if at all, by changing how AI is implemented inside firms, how responsibility for AI-assisted decisions is assigned and how quickly new categories of work absorb the labor that automation displaces. The chapters that follow take up each of these mechanisms in turn, beginning with a closer examination of the claims most often advanced to explain AI's damaged public standing.

2 AI's Negative Impact to Society

The Brookings analysis organizes its case around four principal harms: an erosion of independent thinking through cognitive offloading, job displacement concentrated among routine occupations, a further hollowing out of face-to-face human relationships and the environmental cost of the data centers that power large models. Each deserves to be taken on its own terms rather than dismissed together, because they operate through genuinely different mechanisms and carry different implications for policy. But taken together, they also illustrate a pattern worth naming: several of these harms are real and significant, yet none of them, examined closely, supports the report's ultimate conclusion that better AI policy communication is the missing ingredient in restoring trust.

Cognitive offloading is the most analytically interesting of the four and also the one most often mischaracterized. The underlying phenomenon is well established in cognitive science. Evan Risko and Sam Gilbert's foundational 2016 review in Trends in Cognitive Sciences describes offloading as the use of an external tool or aid to reduce the information-processing demands of a task, such as writing a note instead of memorizing an appointment or using a calculator instead of performing long division by hand.[8] Offloading of this kind is not new and it is not inherently corrosive. The classic 2011 study by Betsy Sparrow, Jenny Liu and Daniel Wegner, published in Science, showed that people readily forget information they expect to be able to look up again, while remembering more reliably where to find it: a strategy that is, within limits, an efficient use of finite attention rather than a failure of it. The concern is not that people delegate cognitive tasks to tools. It is what happens when the tool in question is a generative system whose outputs look authoritative regardless of whether they are correct and whose design does not require the user to retain any structural understanding of the problem being solved.

A useful conceptual distinction, developed recently in the cognitive science literature examining generative AI specifically, separates ordinary cognitive offloading from what researchers have termed cognitive surrender:

the practice of adopting an AI-generated output as one's own conclusion with minimal scrutiny, rather than delegating a narrow subtask while retaining ownership of the reasoning process.[9] A calculator forces the user to know what operation to perform and to recognize an obviously wrong answer; a generative model will produce a fluent, confident, structurally complete answer whether or not it is correct and gives the user little built-in signal that something has gone wrong. Randomized experiments bear this out. A 2026 study using large-scale controlled trials found that access to AI assistance improved task accuracy in the moment but reduced persistence and measurably weakened independent performance once the assistance was withdrawn, with the effect depending heavily on whether users engaged strategically with the tool or delegated to it wholesale.[10] This is precisely the point the Brookings framing understates: cognitive offloading onto AI is troubling not principally because people are choosing to rely on a tool but because a meaningful share of current AI products are engineered in ways that make it unusually easy to stop verifying, without any friction that would prompt a second look. That is a product design failure and an institutional deployment failure, not evidence that offloading itself is the problem and it is a failure with direct consequences in education and early-career training, since students and new workers who never build the underlying judgment have nothing to fall back on once they enter environments where verification is no longer optional. This risk is not merely theoretical: a fall 2025 Pew survey found that 64 percent of American teenagers had used an AI chatbot and about six in ten said that using chatbots to cheat on schoolwork was at least somewhat common at their school, precisely the population for whom offloaded judgment has the most time to compound.

Job displacement is real but its shape is more specific than a general narrative of AI-driven unemployment suggests and the specifics matter enormously for policy. The clearest recent U.S. data, compiled from more than 161,000 job postings by major employers in January 2026 and analyzed by researchers affiliated with Stanford and the AIDE Institute, found that only 13 percent of AI-related postings were aimed at entry-level candidates, while 71 percent targeted senior, experienced workers.[11] This is not evidence that AI eliminates work in aggregate; broader labor-market data from Vanguard, cited in CNN's reporting, found that employment in occupations with high AI exposure actually grew 1.7 percent in the two years after mid-2023, faster than the 1 percent growth in the pre-pandemic period, with real wages in those occupations accelerating from roughly flat growth before the pandemic to 3.8 percent afterward.[12] These occupation-level patterns sit inside a far more uncertain macro forecast: the World Economic Forum projected in 2023 that employers worldwide would eliminate roughly 83 million positions and create about 69 million new ones by 2027, a net contraction of only about 2 percent of current employment, though five-year, economy-wide projections of this kind carry substantially more uncertainty than the occupation-level postings and wage data above.[13] What the postings data shows instead is a narrowing of entry points: firms are willing to pay for workers who can already supervise, correct and integrate AI output but far less willing to invest in training people who cannot yet do so. This is a distinct and more tractable problem than blanket displacement and it points toward a different policy response than either a moratorium on AI adoption or a communications campaign about AI's benefits.

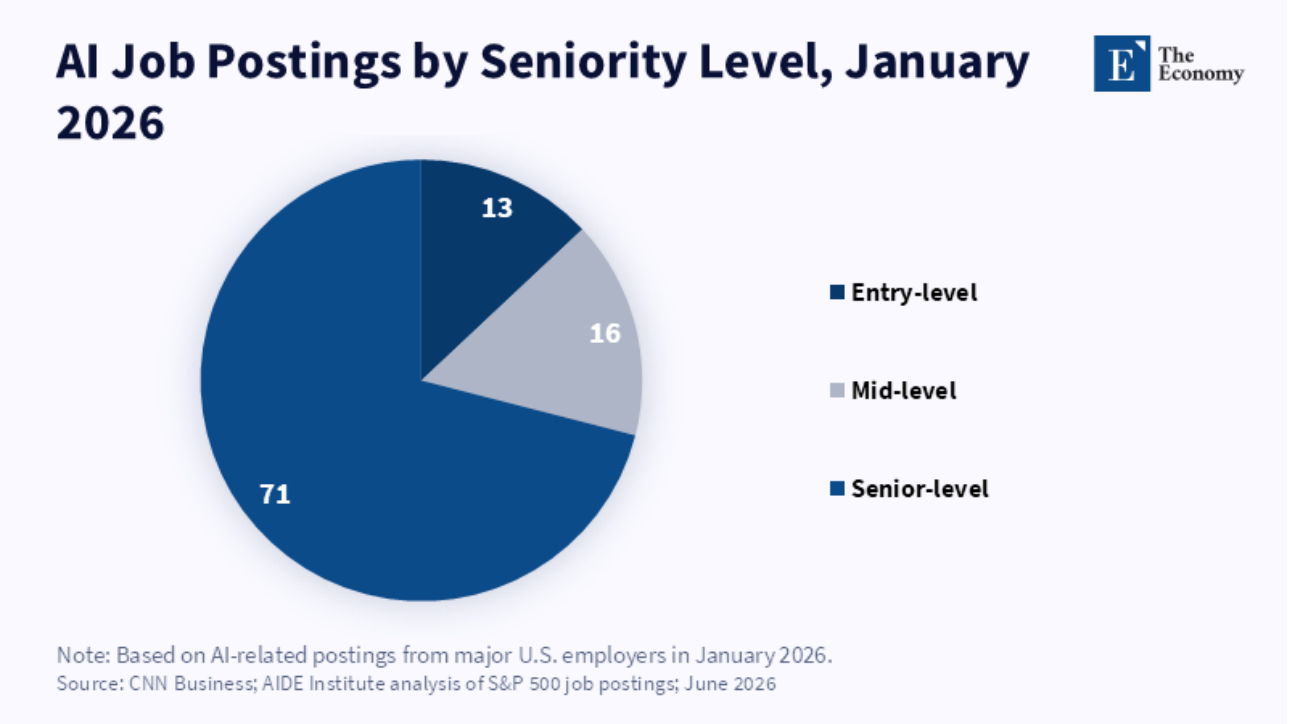

Figure 1: Firms are hiring AI talent almost exclusively at the top, leaving few doors open for beginners.

The claim about human relationships is where the report's own case is weakest and where its author's skepticism is largely justified. It is true that face-to-face social contact has been declining for years, that dating, working and socializing increasingly happen through screens and that AI systems (chat companions, algorithmic matchmaking, remote-work tools) can accelerate an existing trajectory. But the causal engine here long predates generative AI. Remote work scaled sharply during the COVID-19 pandemic; dating shifted onto apps in the 2010s; workplace and social life had already been substantially mediated by digital platforms before large language models existed in their current form. AI intensifies a digitization of social life that was already well underway rather than initiating it and treating it as a freestanding cause of relational decline both overstates its explanatory power and understates the role of the broader platform economy that preceded it by more than a decade.

The environmental argument is, on the evidence, the weakest link connecting stated concerns to AI's actual reputational problem among the public, Gen Z included. Data center energy and water use are legitimate objects of environmental policy and the Brookings report is right that communities near new AI infrastructure deserve disclosure standards and meaningful consultation. But there is little evidence that environmental cost is what drives an ordinary user's frustration with an AI system and considerably more evidence, as the next chapter shows, that the frustration is generated directly by the interaction itself, by AI misunderstanding a claim, looping a customer through the same unhelpful answer or blocking access to a human being who could actually resolve the problem. A policy framework organized around data center siting will do little to change how the technology is experienced at the point of use, which is where distrust is actually being manufactured.

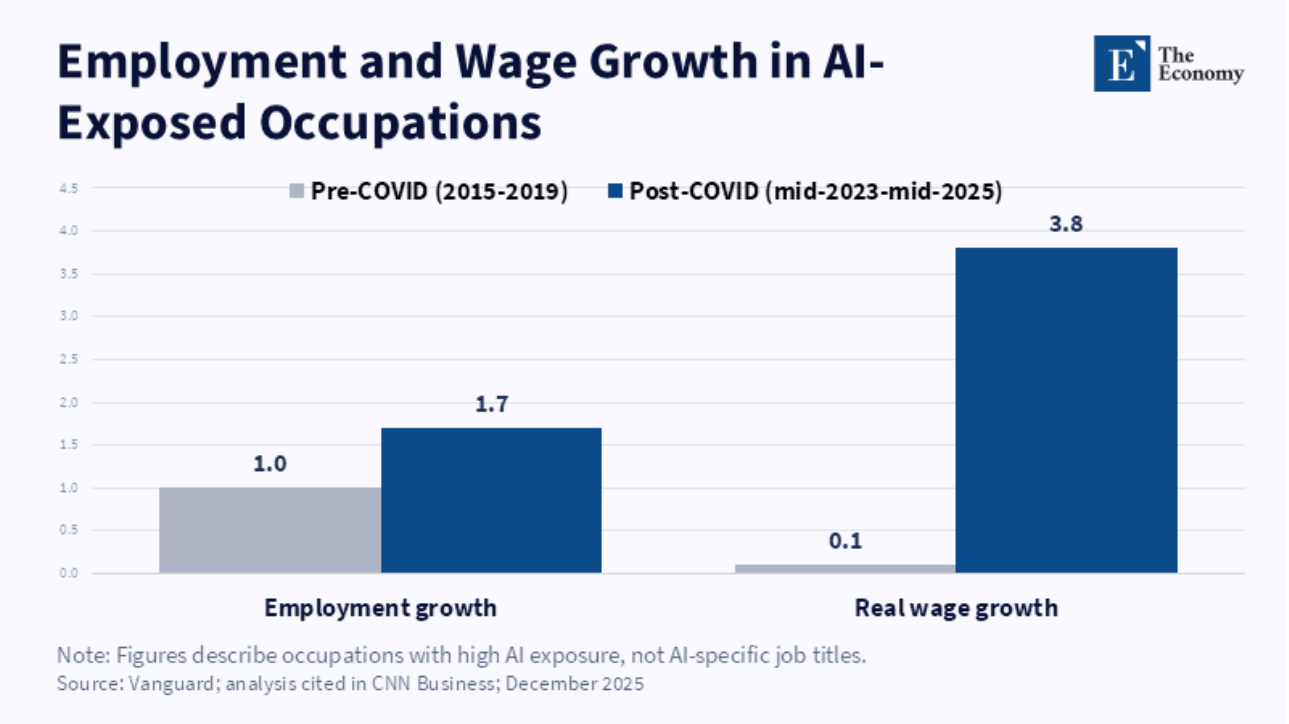

Figure 2: Jobs most exposed to AI are growing faster and paying more since the pandemic, not disappearing.

3 What's Actually Driving the Negative Image?

If the explanation is not primarily environmental cost or a failure of corporate messaging, the more direct account is that AI's negative image is being generated continuously, in millions of small transactions, by the gap between what deployed systems promise and what they deliver. Two distinct mechanisms are doing most of the work here and it is worth separating them cleanly, because they call for different remedies. One is a limitation intrinsic to the technology itself: current models are probabilistic, can produce plausible but incorrect output and are not yet reliable enough to be trusted without oversight in many of the contexts where firms have deployed them. The other is a limitation of implementation: firms have often deployed AI in ways optimized for cost reduction rather than for the customer's actual problem, stripping out the human escalation paths that used to catch AI's failures before they reached the customer.

The implementation failure is documented with unusual clarity in customer service research. A practical analysis of AI support frustration identifies what its authors call the infinite loop problem: a customer's issue goes unresolved, the system repeats the same unhelpful suggestion and escalation to a human becomes harder rather than easier precisely when the customer is most frustrated.[14] A separate and more consequential failure is the loss of context at the point of handoff: when a customer does eventually reach a human agent, they are frequently asked to restart the entire explanation from the beginning, because the AI system's understanding of the problem is not carried forward. Academic research on service perception gives this pattern a theoretical foundation. A 2025 study published in the Journal of Retailing and Consumer Services, based on five experiments plus a supplemental study, found that replacing human staff with AI service agents significantly reduced customers' perceived service warmth and traced this effect to diminished perceptions of the AI agent's experience and agency, the qualities customers unconsciously attribute to a mind capable of understanding their

situation.[15] Warmth, in this research tradition, is not a soft or marginal variable; it is closely tied to trust, continued usage intention and willingness to forgive a service failure. When a firm replaces a human agent with an AI system that customers do not perceive as capable of understanding them, it is not simply changing the channel of service delivery. It is removing a psychological ingredient that made customers tolerant of imperfection in the first place and that removal shows up directly in review scores and repeat-business behavior.

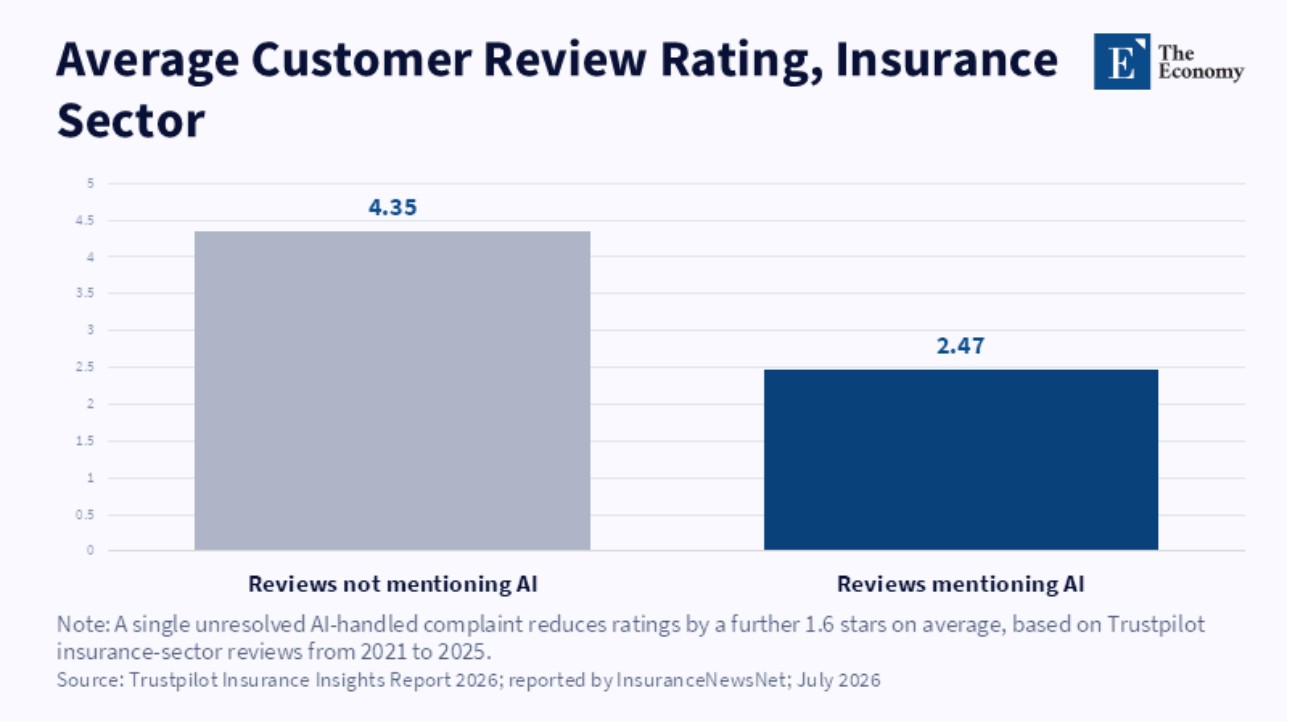

Figure 3: Reviews that mention AI service score nearly two full stars lower than those that don't.

None of this means firms were wrong to experiment with AI-driven service or that automation has no place in customer interaction. The Trustpilot findings referenced earlier are instructive on this point precisely because they are not a blanket indictment of automation: consumers appear willing to engage with AI when it demonstrably makes an interaction faster but the tolerance evaporates when AI becomes, in the words of Trustpilot's own analysis, a barrier that keeps customers from reaching a human being who can exercise judgment. The distinction being drawn by consumers, whether or not they would phrase it this way, tracks closely onto the difference between assistance and replacement. AI that assists a human agent (summarizing a case history, drafting a response the agent reviews, surfacing relevant policy details) preserves the accountability structure that a service relationship depends on. AI that replaces the human agent outright, particularly in emotionally significant or high-stakes interactions, removes that structure and customers notice the removal even when they cannot articulate the theory of mind research explaining why.

The market itself has begun to register this distinction, which is telling, because market corrections tend to follow revealed costs rather than public relations narratives. Gartner's Customer Service and Support practice, in a forecast released in February 2026, projected that by 2027, half of the companies that attributed customer service headcount reductions to AI will rehire staff to perform similar functions, in many cases simply under different job titles. Gartner's own analysts framed the underlying cause plainly: current AI is not mature enough to replace the expertise, empathy and judgment that human agents supply in complex or emotionally

charged cases and firms that cut too aggressively are now confronting the operational cost of that miscalculation, overloaded remaining staff, degraded service quality and eroded customer trust that is expensive to rebuild.[16] That correction follows an earlier Gartner admission: a June 2025 report found that only about a fifth of customer service leaders had actually reduced agent staffing because of AI, even as a March 2025 poll of 163 service and support leaders found 95 percent already planning to retain human agents as a matter of strategy.[17] This is a market signal, not a policy signal and it points toward a conclusion that a regulatory framework governing AI marketing claims would not meaningfully affect: firms that deployed AI as a headcount-reduction tool rather than a capability-augmentation tool are now paying for that choice in customer attrition and service failure and are correcting course because the correction is cheaper than continuing to absorb the damage.

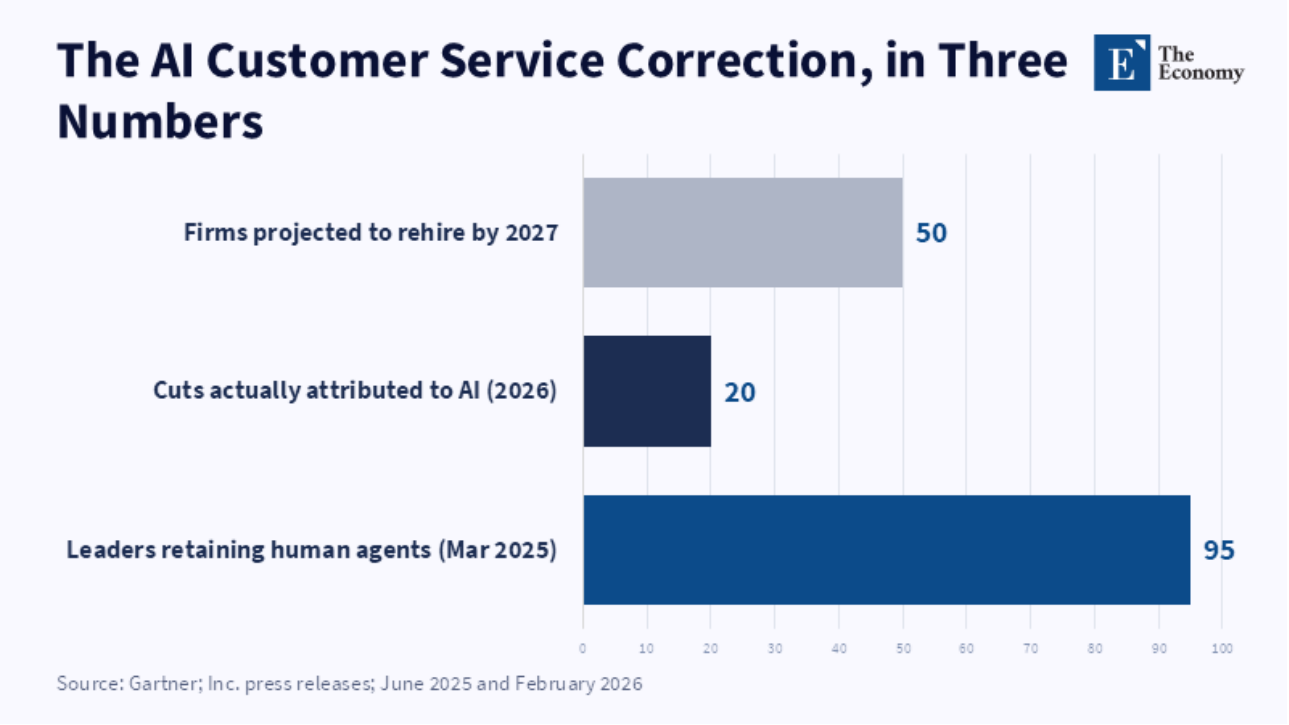

Figure 4: Only a fifth of recent cuts were really AI-driven, yet half of firms are already reversing course.

A further distinction sharpens why the implementation failure is the more tractable of the two problems identified here. Producing an output and accepting responsibility for it are not the same act and a great deal of what customers experience as AI's untrustworthiness traces back to firms collapsing that distinction rather than to the underlying model being unusually unreliable. When a human claims adjuster tells a policyholder that a claim is denied, the adjuster is a locatable person who can be asked to explain the reasoning, escalate the decision or be held accountable if the reasoning was wrong. When an AI system delivers the same denial without a clear, reachable path to a person who owns that decision, the customer is not simply receiving a different channel of the same information; they are receiving a determination that nobody in the organization has visibly agreed to stand behind. This is the practical difference between assistance, where AI supports a human who retains ownership of the outcome and reliance shading into over-reliance, where the firm has quietly transferred ownership to a system incapable of being held accountable in any meaningful sense. Firms that have preserved this ownership structure, in which AI drafts and a person decides, report far less of the reputational damage documented in the Trustpilot data, because the customer can still reach the person who is answerable for the

result. Firms that removed the human from the loop entirely, in the name of efficiency, are the ones now facing the rehiring correction Gartner has projected. The lesson is not that AI cannot participate in consequential decisions; it is that participation and ownership need to remain visibly distinct and that the erosion of trust tracks the disappearance of an accountable person far more closely than it tracks any property of the underlying model.

It is worth being precise about what the program-level limitation actually consists of, because conflating it with the implementation failure obscures where responsibility lies. Even well-implemented AI systems, with clear escalation paths and preserved context, remain probabilistic tools that can generate confidently wrong output, a property sometimes called hallucination, though the term understates how ordinary the failure mode is in practice. Pew's own polling captures the downstream consequence among the minority of Americans who already rely on AI chatbots for information: about half of those who get news from AI chatbots report encountering content they believe to be inaccurate at least sometimes, including 16 percent who say this happens often or extremely often. This is not a problem a disclosure policy or an advertising standard can fully solve, because the failure is not that companies are overpromising in their marketing; the failure is a property of the underlying technology that persists regardless of what any company claims about it. What good implementation can do (preserving human escalation, using AI to augment rather than replace judgment and being honest with users about confidence levels) is bring the deployed system's actual reliability closer to what a well-informed user would reasonably expect. What implementation cannot do, at least with the current generation of models, is eliminate the underlying error rate. Distinguishing these two problems clarifies why the policy response belongs substantially inside firms and professional bodies that govern deployment standards, rather than in a public messaging campaign directed at end users.

4 What We Need: AI to Create New Adaptation and Thus New Jobs

If distrust is generated by the gap between what AI promises and what it reliably delivers and by firms deploying it in ways that strip out the human judgment customers still require, the more durable remedy is not a better explanation of AI's benefits but a labor market that visibly absorbs displaced work into new, better-compensated roles. Once AI participation becomes a demonstrated source of career advancement rather than a source of career risk, the negative public image identified in the preceding chapters becomes a secondary concern, because people's felt experience of the technology shifts from something done to them toward something that works for them.

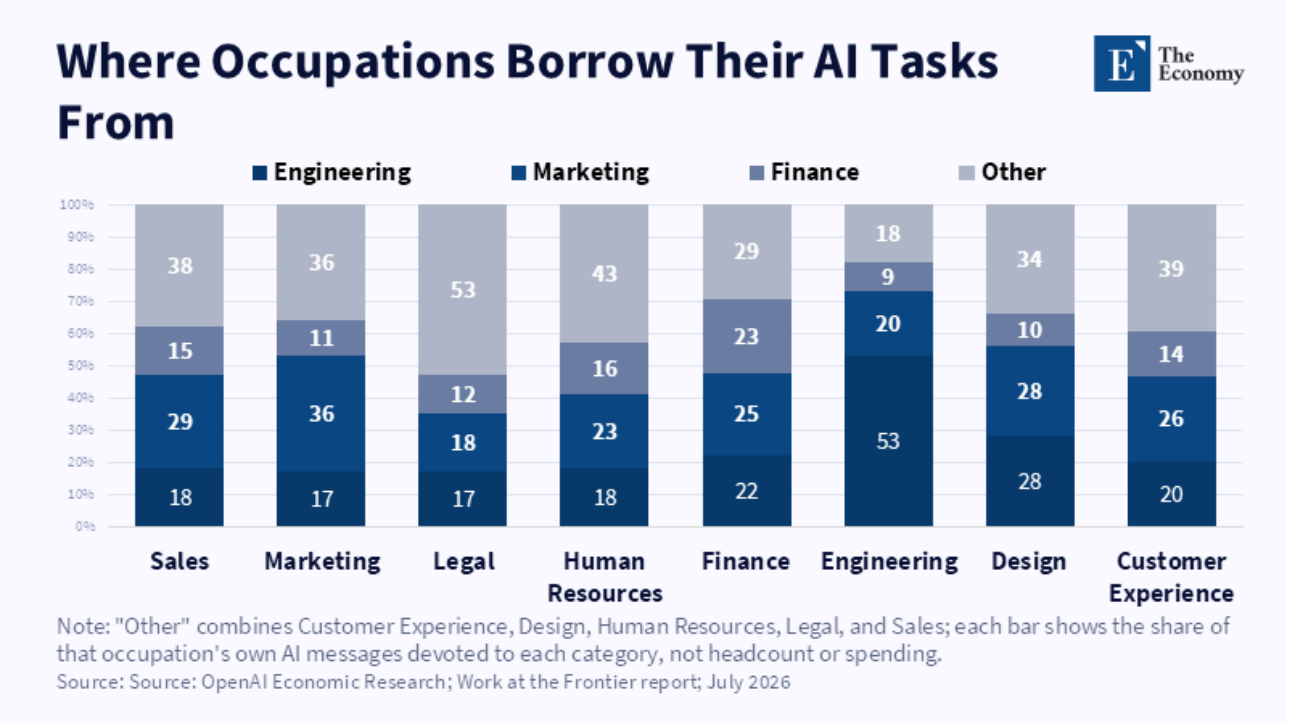

Figure 5: Nearly every occupation now relies on marketing and engineering tasks outside its own domain.

The evidence that this absorption is already occurring, unevenly but measurably, is stronger than the more alarmist commentary about AI-driven unemployment usually acknowledges. This kind of shift extends a task-based view of technological change with roots in labor economics, which shifted attention away from occupations as fixed units and toward the discrete activities that workers and machines perform,[18] and which later emphasized that automation can both reallocate existing tasks to capital and create new tasks in which labor retains a comparative advantage.[19] OpenAI's Economic Research team, analyzing more than 800,000 U.S. ChatGPT messages in a July 2026 report, found that 43.5 percent of occupation-specific AI use involved tasks conventionally associated with a different occupation than the user's own, a pattern the researchers term task crossover. A small-business owner drafting a contract that would once have gone to outside counsel, a salesperson exploring a dataset that would once have required an analyst, a marketer troubleshooting a website without waiting on a developer: in each case, AI is not eliminating the underlying work so much as redistributing who is capable of doing it. This crossover was especially pronounced among customer experience workers, 77 percent of whose occupation-specific messages involved tasks outside their traditional role and among designers, at 75 percent. Across all eight occupation groups the researchers studied, this share ranges from 28 percent to 77 percent and it is somewhat more pronounced at smaller firms: among typical-volume users, the cross-occupation share runs 18.9 percent in workspaces of two to five seats against 16.3 percent in workspaces of 101 or more, a gap consistent with the small-business pattern described above, where fewer colleagues and specialists are available to hand a task to. The task list itself, in the researchers' words, is changing before job titles or formal descriptions catch up, which means official labor statistics, tied to fixed occupational categories, are likely to register this shift only with a lag, understating how much genuine adaptation is already underway.

This pattern is consistent with longer-running economic research on how technology reshapes rather than simply eliminates employment. A March 2025 study using instrumental-variable methods to separate the labor-

market effects of automating AI, which substitutes for workers, from augmenting AI, which enhances what workers can do, found that augmentation AI is associated with the emergence of new job titles and rising wages, concentrated among higher-skilled occupations, while automation AI's negative effects fall disproportionately on lower-skilled work.[20] The distributional consequence is real and should not be minimized: technology of this kind does not create equally distributed benefits and a rising wage premium concentrated among workers who can already direct and evaluate AI output is a mechanism that widens inequality rather than narrowing it. But the finding also confirms that new work formation is an empirically observed response to AI adoption, not merely an aspirational talking point. This matters for policy design, because it means the correct target is not whether new jobs will appear, the evidence suggests many already are but who has access to the training and entry pathways that let them move into those jobs before the gap between AI-fluent and AI-naive workers hardens into a durable inequality.

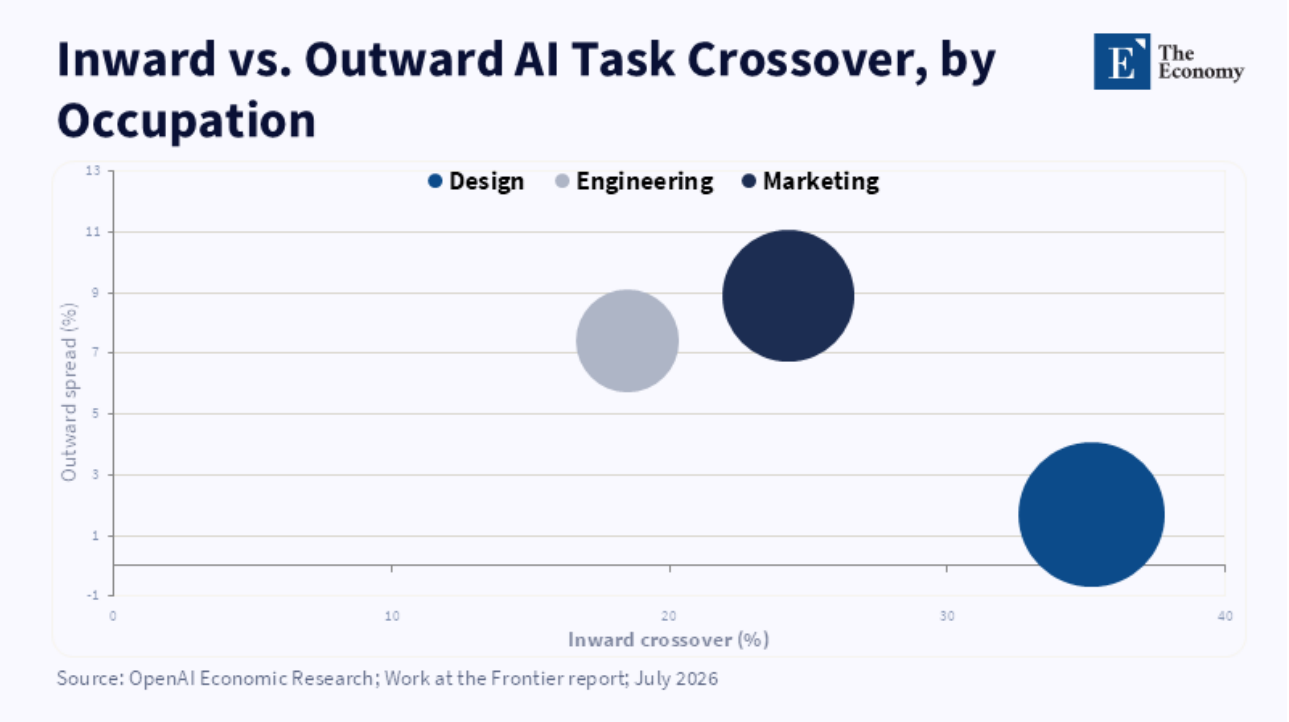

Figure 6: Designers borrow far more than they lend; engineers mostly export tasks.

The software engineering labor market illustrates both the reorganization and its limits with unusual clarity, precisely because it is the occupation most exposed to AI capability. Job cuts attributed to AI are real and documented: outplacement firm Challenger, Gray & Christmas recorded more than 49,000 job losses citing AI as a factor in a recent year and companies including Block and Coinbase have explicitly linked staff reductions to AI-driven productivity gains, with Block cutting roughly 40 percent of its workforce and Coinbase around 14 percent.[21] At the same time, this is best read as a reorganization of the profession rather than its disappearance. As Anthropic's own head of Claude Code told CNN in 2026, the substance of the job is shifting away from writing lines of code and toward reviewing, architecting and directing AI-assisted output, the same distinction, expressed in labor-market terms, between offloading a subtask and surrendering the underlying judgment that Chapter 1 identified in the cognitive-science literature. A separate analysis tracking AI spending and employment data across roughly 22,000 U.S. firms found that companies investing in AI as a strategic capability, rather

than treating it purely as a cost-cutting measure, showed the strongest employment momentum, though the researchers cautioned that firms making the heaviest AI investments were already larger, more capital-intensive and faster-growing before adopting the technology, a reminder that correlation between AI spending and job growth should not be read uncritically as AI's causal effect.[22] The overall picture that emerges from this evidence is neither the wholesale job destruction implied by the most alarmed commentary, nor the frictionless job creation implied by the most enthusiastic industry voices but something more specific: a reorganization that rewards workers who can supervise and direct AI, penalizes routine execution that AI can already perform reliably and currently underserves the entry-level workers who would otherwise grow into the supervisory roles the market is creating.

This is precisely where the Brookings framing, for all its attention to labor protections, arrives at a policy conclusion that undersells what is actually needed. Portable benefits, wage insurance and disclosure standards for AI-attributed layoffs are reasonable safety-net measures and there is a legitimate role for government in providing them, particularly given that individual firms have limited incentive to fund transition costs for workers they are laying off. But safety-net policy treats job loss as the terminal event to be cushioned, rather than treating the absence of accessible entry pathways into AI-augmented roles as the actual structural problem to be solved. The AIDE Institute data on the seniority skew of AI job postings points to a more targeted intervention: apprenticeship-style structures, explicit junior-track hiring requirements attached to public contracts and employer tax incentives tied to entry-level AI-skills training would address the mechanism directly, rather than compensating workers after the fact for a labor market that never let them in. Firms, for their part, bear direct responsibility here that public policy cannot substitute for: the CEO of the AIDE Institute, in comments accompanying the January 2026 postings data, argued that corporate leaders need to prioritize AI talent development at every level of the organization, not only at the senior level where the current hiring pattern is concentrated, because a firm that trains no junior AI talent today is simply exporting its future talent shortage to competitors and public institutions. That is a governance failure inside firms and no plausible AI marketing regulation addresses it.

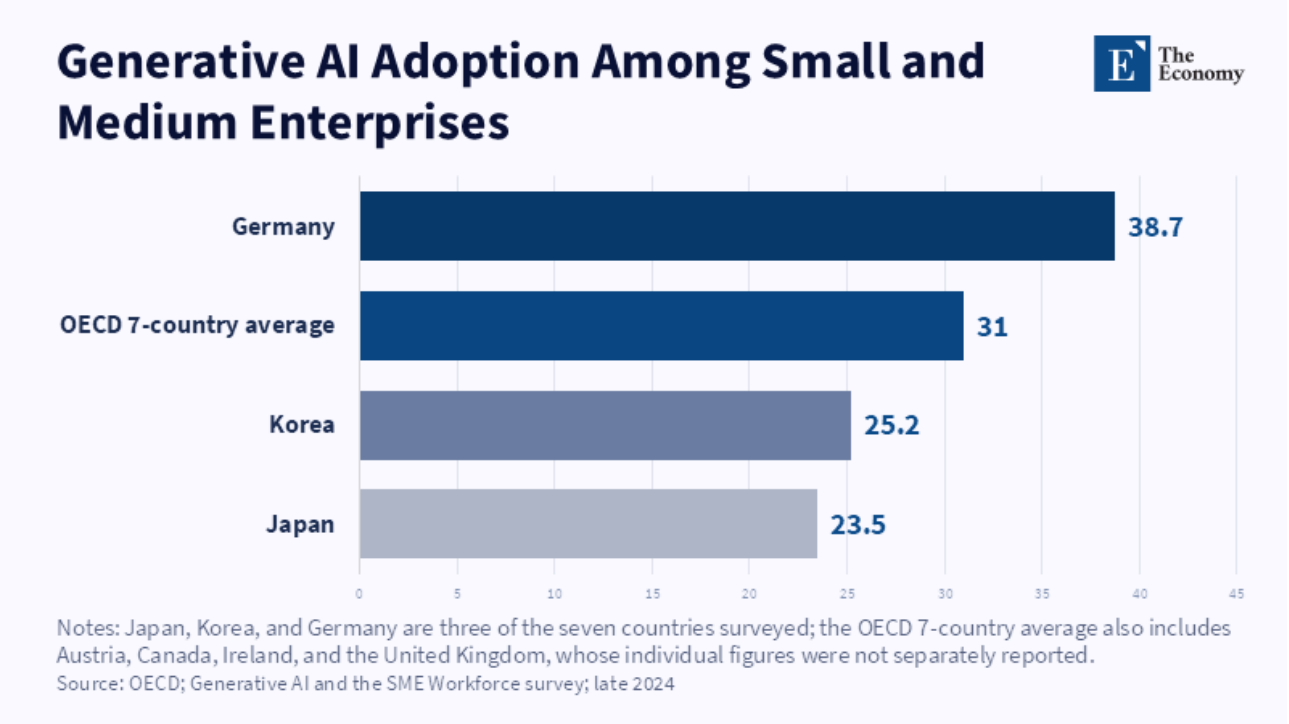

Figure 7: Younger employees already lead workplace AI use, even though hiring still rewards experience.

A related distinction between the adoption of AI and the governance of AI helps explain why the labor-market reorganization documented above is proceeding faster than institutions' capacity to manage it fairly. Adoption is now essentially universal at the firm level; a representative OECD survey of more than 5,000 small and medium enterprises across seven countries found generative AI already in use at 31 percent of firms on average by late 2024, ranging from 23.5 percent in Japan and 25.2 percent in Korea up to 38.7 percent in Germany, figures that have almost certainly risen further since.[23] Governance, meaning deliberate firm-level decisions about which roles get AI-augmentation training, which tasks remain human-only and how newly automatable work is redistributed rather than simply eliminated, has not kept pace with adoption in most organizations. The same OECD survey found that fewer than a third of SMEs already using generative AI reported that their employees received any related training, a share that fell as low as 11.3 percent in Japan, even though half of surveyed SMEs said a lack of AI skills among staff was already holding the technology back. This is precisely the combination of skills that entry-level workers have historically developed on the job, through supervised practice on the routine tasks that AI now performs instead, which means the seniority skew documented in the AIDE Institute postings data is not simply a hiring preference; it reflects the removal of the very training ground junior workers once used to acquire the judgment that senior roles now require. Governing AI adoption well, in this sense, means firms deliberately reconstructing that training ground rather than assuming it will persist on its own once the tasks that used to constitute it have been automated away.

None of this amounts to uncritical technological optimism and the strongest version of the counterargument deserves to be stated in its own terms before being qualified. AI genuinely does raise productivity in measurable ways, genuinely does extend access to capabilities (legal drafting, financial analysis, basic design) that were previously gated behind specialist training and genuinely does remove some of the most repetitive components of many jobs, in a lineage that traces back through calculators, spreadsheets and search engines rather than

representing something categorically unprecedented. Economic historians have found that roughly 60 percent of U.S. employment in 2018 existed in job titles that did not exist in 1940, evidence that new work formation in response to technological change is itself a long-standing pattern rather than a novel promise.[24] Vanguard's finding of accelerated employment and wage growth in AI-exposed occupations supports this case directly, as does the OECD's assessment that AI's net employment effect depends heavily on whether it functions as an automating or augmenting force in a given task.[25] The honest qualification is not that this optimistic account is wrong but that it is incomplete in a specific and correctable way: the benefits it describes accrue disproportionately to workers who already possess the judgment to direct AI output, while the costs fall disproportionately on workers who have not yet had the chance to acquire that judgment and the current market, left alone, is not closing that gap on its own. Policy and firm practice that target the gap directly, rather than either restraining AI adoption or promoting it rhetorically, are what would actually shift public sentiment, because sentiment is downstream of lived experience and lived experience is currently telling a large number of workers, especially younger ones, that AI's benefits are real but are not yet reliably available to them.

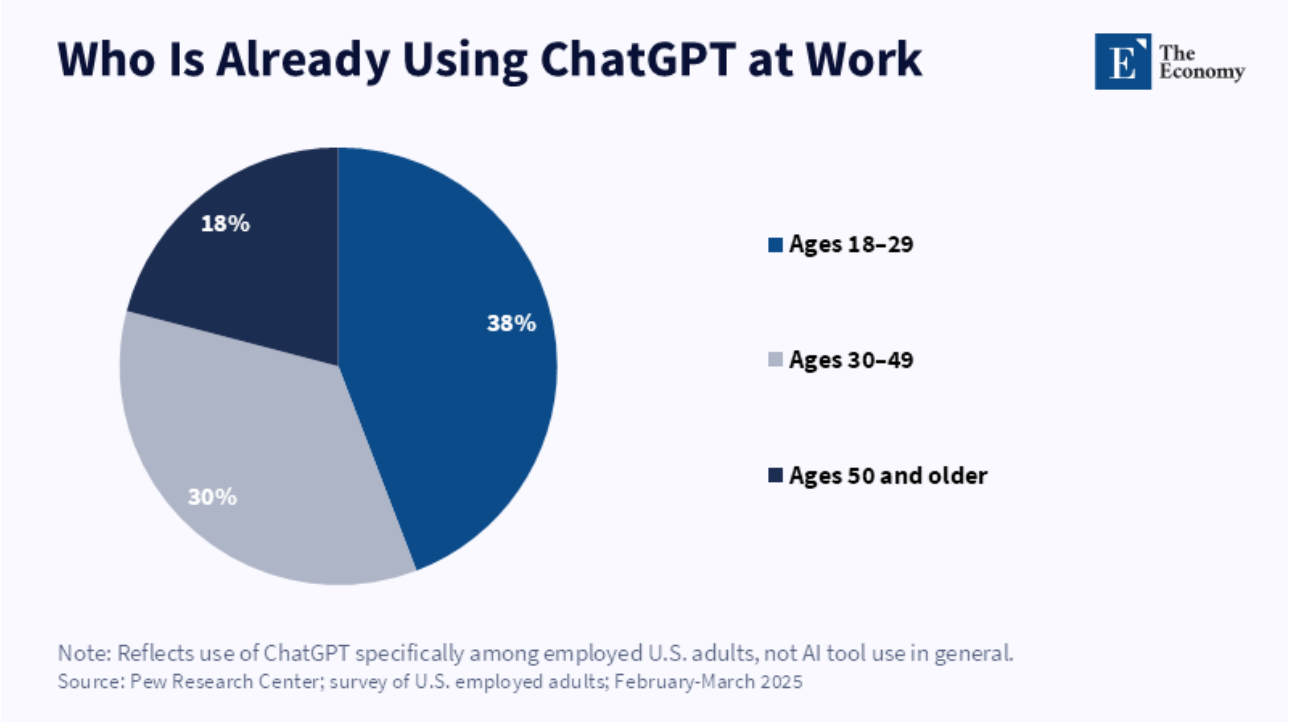

Figure 8: SME adoption already spans a 15-point range across peer economies surveyed in the same wave.

5 Conclusion - What Actually Rebuilds Trust in AI

The distrust surrounding artificial intelligence is not primarily a communications problem and it will not be resolved by advertising standards or more disciplined corporate messaging. It is generated by a specific set of mechanisms: systems that make cognitive surrender easier than verification, customer-facing deployments optimized for cost reduction rather than for the judgment customers still need and a labor market reorganizing around AI in ways that so far reward mainly workers who already possess the skills to supervise it. Each mechanism produces a measurable cost (lower service ratings, weakened independent performance once AI assistance is withdrawn, a hiring market skewed toward the already experienced) and each is addressable closer

to where the harm originates than through rules governing what AI companies may claim about their products.

The correction already visible in customer service, where roughly half of companies that cut staff for AI are projected to rehire by 2027, shows that markets absorb some of this mismatch on their own, though slowly and at real cost in the interim. Public policy has a legitimate but narrower role than the communications-focused framing suggests: funding training pathways into AI-supervisory roles, requiring disclosure when AI substantially drives a layoff and setting escalation standards for consequential AI-mediated services, rather than regulating marketing claims that were never the primary driver of public sentiment. Firms bear a correspondingly direct responsibility to treat AI as an augmentation of judgment rather than its substitute and to build junior-level AI fluency now rather than treat the current seniority skew as someone else's problem. What ultimately determines whether AI is trusted is not what companies say about it but what happens the next time a customer files a claim, a graduate applies for a first job or a worker asks whether the tool she uses daily will still have a place for her in five years. Policy and firm practice that answer those questions directly will do more to restore trust than any campaign built around what AI companies choose to advertise.

References

[1] Morgan, R. (2026) 'AI communication gap' leaves negative impression on customers. *InsuranceNewsNet*, 3 July.

[2] Stewart, J. and Tanner, B. (2026) 'Policy, not PR, will determine Gen Z's trust in AI'. *Brookings*, 22 July.

[3] Pew Research Center (2026) *Americans' Views on AI Chatbots, Smart Devices and AI's Impact.* Washington, DC: Pew Research Center, 17 June.

[4] Faverio, M. and Kikuchi, E. (2026) *Key Findings About How Americans View Artificial Intelligence.* Washington, DC: Pew Research Center, 12 March.

[5] Gartner, Inc. (2026) *Gartner Predicts Half of Companies That Cut Customer Service Staff Due to AI Will Rehire by 2027.* Press release, 3 February.

[6] Risko, E.F. and Gilbert, S.J. (2016) 'Cognitive offloading', *Trends in Cognitive Sciences*, 20(9), pp. 676–688.

[7] OpenAI (2026) *How AI Is Expanding What People Do at Work.* San Francisco: OpenAI Economic Research, 27 July.

[8] Sparrow, B., Liu, J. and Wegner, D.M. (2011) 'Google effects on memory: cognitive consequences of having information at our fingertips', *Science*, 333(6043), pp. 776–778.

[9] Shaw, J. and Nave, G. (2026) 'Cognitive surrender: distinguishing generative-AI reliance from cognitive offloading', as discussed in *Faster Completion, Less Learning: Generative AI Reduced Study Time on Math Problems and the Knowledge They Build.* arXiv preprint.

[10] *AI Assistance Reduces Persistence and Hurts Independent Performance* (2026) arXiv preprint.

[11] Duffy, C. (2026) 'AI is sparking a jobs boom: just not for newbies', *CNN Business*, 11 June.

[12] *The Surprising Truth About AI's Impact on Jobs* (2025) CNN Business, 18 December.

[13] World Economic Forum (2023) *The Future of Jobs Report 2023*. Geneva: World Economic Forum.

[14] Gleap Team (2026) *Why Customers Are Frustrated by AI Customer Service, and How to Fix It in 2026*. Gleap, 27 April.

[15] Lv, X., Yang, Y., Qin, D. and Liu, X. (2025) 'AI service may backfire: reduced service warmth due to service provider transformation', *Journal of Retailing and Consumer Services*, 85.

[16] Nguyen, T. (2026) 'Why AI-first customer service strategies are about to backfire', *CMSWire*, 23 February.

[17] Gartner, Inc. (2025) *Gartner Predicts 50% of Organizations Will Abandon Plans to Reduce Customer Service Workforce Due to AI*. Press release, 10 June.

[18] Autor, D.H., Levy, F. and Murnane, R.J. (2003) 'The skill content of recent technological change: an empirical exploration', *Quarterly Journal of Economics*, 118(4), pp. 1279–1333.

[19] Acemoglu, D. and Restrepo, P. (2019) 'Automation and new tasks: how technology displaces and reinstates labor', *Journal of Economic Perspectives*, 33(2), pp. 3–30.

[20] Marguerit, D. (2025) *Augmenting or Automating Labor? The Effect of AI Development on New Work, Employment, and Wages*. LISER Working Paper. arXiv preprint.

[21] *AI Isn't Actually 'Taking' Your Job. Here's What's Happening Instead* (2026) CNN Business, 10 May.

[22] *Study: AI Is Actually Creating More Jobs, Not Killing Them (But There's a Catch)* (2026) 247wallst/Yahoo Finance, citing Ramp Economics Lab and Revelio Labs research reported by the Financial Times.

[23] OECD (2025) *Generative AI and the SME Workforce*. Paris: OECD Publishing.

[24] Autor, D., Chin, C., Salomons, A. and Seegmiller, B. (2024) *New Frontiers: The Origins and Content of New Work, 1940–2018*. NBER Working Paper.

[25] OECD (2026) *The Impact of Artificial Intelligence on the Labour Market*. Paris: OECD Publishing.