The Future of Work in the Age of AI: Displacement or Risk-Shifting?

Pegah Moradi and Karen Levy
The Oxford Handbook of Ethics of AI
Edited by Markus D. Dubber, Frank Pasquale, and Sunit Das

Abstract and Keywords

This chapter examines the effects of artificial intelligence (AI) on work and workers. As AI-driven technologies are increasingly integrated into workplaces and labor processes, many have expressed worry about the widespread displacement of human workers. The chapter presents a more nuanced view of the common rhetoric that robots will take over people’s jobs. We contend that economic forecasts of massive AI-induced job loss are of limited practical utility, as they tend to focus solely on technical aspects of task execution, while neglecting broader contextual inquiry about the social components of work, organizational structures, and cross-industry effects. The chapter then considers how AI might impact workers through modes other than displacement. We highlight four mechanisms through which firms are beginning to use AI-driven tools to reallocate risks from themselves to workers: algorithmic scheduling, task redefinition, loss and fraud prediction, and incentivization of productivity. We then explore potential policy responses to both displacement and risk-shifting concerns.

Keywords: artificial intelligence, AI-driven technologies, workplaces, labor processes, displacement, human workers, AI-induced job loss, task execution, algorithmic scheduling, fraud prediction

IN February 2011, Jeopardy! viewers watched as the AI system known as IBM Watson defeated Ken Jennings and Brad Rutter, two of the winningest Jeopardy! champions of all time, in a three-day exhibition match The New York Times lauded as “a vindication for the academic field of artificial intelligence.”¹ Watson’s ability to understand and respond to Jeopardy! clues was considered a major step forward for natural language processing and information retrieval, and soon after, IBM announced plans to use the system to assist physicians in making diagnoses or treating patients.²

Winning at Jeopardy! was a unique challenge for a machine, given that Jeopardy! is more unpredictable and complex than a simple test of trivia; as Jennings wrote in 2019, its clues are “weird, short little haikus, laced with hints, puns, winks, and red herrings.”³ When Watson erred, it often seemed to miss clues that humans would find easy or obvious. Watson, for example, rendered “what is chic?” in response to the clue “stylish ele-
The Future of Work in the Age of AI: Displacement or Risk-Shifting?

gance, or students who all graduated in the same year”; Brad Rutter subsequently offered the correct response, “what is class?” In a Final Jeopardy! round with the category “U.S. Cities,” Watson responded, “What is Toronto????” with four question marks denoting low confidence in the response.

But despite its shortcomings, Watson still won. Many assumed that this was simply because Watson had a memory capacity of fifteen trillion bytes and had been fed data from millions of documents, books, encyclopedias, and news articles. Watson was able to consume a wealth of information that most people—even Jeopardy! champions—could only dream of being able to absorb. But it is also possible that a much simpler mechanism gave Watson the biggest advantage of all: Jennings suggests that Watson was so good largely because it was much quicker to the buzzer than its human competitors were. “As Jeopardy! devotees know,” Jennings notes, “if you’re trying to win on the show, the buzzer is all. On any given night, nearly all the contestants know nearly all the answers, so it’s just a matter of who masters buzzer rhythm the best.” In response to criticism over Watson’s buzzer advantage, IBM researcher Eric Brown noted: “there are some things that computers are going to be better at than humans and vice versa. Humans are much better at understanding natural language. Computers are better at responding to signals.”

The combination of comparative strengths and weaknesses that Watson brought to the Jeopardy! stage nicely encapsulates the nuanced relationship between AI and human work. The computer’s success was seen as a bellwether, as futurists used Watson’s win as a launch pad for claims about the possibility of AI displacing workers. (“After all,” fretted Martin Ford, “if a machine can beat humans at Jeopardy!, will computers soon be competing with people for knowledge-based jobs?”) In some respects, Watson’s abilities were far superior to those of its human competitors—but humans were innately capable of aspects of gameplay with which Watson struggled. Though the specifics of the task may differ, the same is true of all human/machine relations in work contexts.

To understand the ethical issues most likely to beset the future of work, we must first realistically assess what kinds of threats AI might pose. Though some economists and policymakers have begun to express great concern about what AI will mean for employment—including whether some forms of work will exist at all—we argue that the popular “robots will take our jobs!” narrative of AI-induced job displacement is overly simplistic and alarmist. In spite of rapid growth in research and in application, AI systems still have quite limited practical capabilities, and the current technical limitations of AI still give humans the comparative advantage in many kinds of work. Forecasts of widespread employment displacement tend to focus solely on technical aspects of work, and neglect broader contextual inquiry about the social components of work, organizational structures, and cross-industry effects. In the first part of this chapter, we explain these limitations of existing forecasts.

In the second part, we turn to the outcomes we do expect from AI in the workplace. Specifically, intelligent systems are likely to be marshaled toward traditional man-
The Future of Work in the Age of AI: Displacement or Risk-Shifting?

agerial goals related to efficiency, productivity, and risk mitigation. We highlight four ways in which firms may use AI in pursuit of these goals, effectively offsetting risks from themselves onto their workers. We end with discussion of potential policy responses to these concerns.

AI as Worker Displacement: Rhetoric and Reality

As AI-driven technologies are increasingly integrated into work processes, a commonly expressed concern is the impending displacement of human workers—often apocalyptically phrased in popular media as “robots taking over our jobs.” This argument tends to follow from the understanding that human work is comprised of a series of tasks, some or all of which can be done more effectively, efficiently, or at scale by a machine. Therefore, as machines grow in capability, a greater number of tasks currently performed by humans can (and, it is assumed, will) be automated. Because human work is comprised of these tasks, the thinking goes, human workers are vulnerable to being displaced by machines—potentially leaving many without jobs or drastically rearranging how labor is distributed by occupation. And because the jobs widely believed to be most acutely threatened by AI are blue-collar jobs—often held by less educated and poorer workers with fewer alternative options—there is, it is feared, potential for tremendous social and economic disruption.

What Kinds of Tasks Can AI Execute?

Machines are newly capable of performing a number of tasks formerly “off limits” to automation, thanks to technical improvements in AI, increased access to big datasets, and advancements in robotics. Prior to these developments, the paradigmatic model of task-based automation was the two-factor model proposed by Autor, Levy, & Murnane in 2003, which we will refer to as the ALM model. ALM focuses on how routine a task is on one dimension, and the degree to which tasks involve cognitive versus physical work on the other dimension. As Autor and his co-authors argued, “computer capital” could substitute for workers executing abstractable, programmable routine tasks—consisting of both “cognitive and manual tasks that can be accomplished by following explicit rules.” Watson’s buzzer advantage was rooted in this specific routine capability: being able to respond quickly and predictably to an explicit signal. The ALM model posited that nonroutine human labor might be complemented by computers, but that computers were unlikely to substitute wholly for humans for nonroutine tasks. Nonroutine tasks were deemed more difficult to program and dependent on skills like perception, problem-solving, and intuition that were well beyond the purview of computing in 2003.

But the world has changed since then. As computers have become more sophisticated and responsive to their environments, they can adapt to dynamic situations more adeptly—negotiating traffic, responding to conversational cues, developing novel solutions to problems. In light of robotic capabilities, computer vision, and machine learning, it’s less
The Future of Work in the Age of AI: Displacement or Risk-Shifting?

important than it once was that a task be clearly definable and repeatable, thus complicating the ALM model. With AI, many tasks previously thought to be intractably nonroutine are becoming converted into abstractable problems aided by the availability of large and complex datasets.\textsuperscript{12} Although machines were previously limited to tasks that were clearly defined with limited potential contingencies, today’s AI systems can analyze previous cases to determine a course of action in unpredictable situations. Likewise, integrating prediction-driven models with robotics can bring these capabilities into the realm of physical labor. For instance, though Autor et al. explicitly mentioned truck driving as a manual nonroutine task in their 2003 work (and hence likely to be safe from automation), several companies have set goals to develop fully autonomous long-haul vehicles in the near future based on new technical capabilities.\textsuperscript{13}

While AI can allow a machine to execute tasks that would have previously been considered nonautomatable under the ALM model, AI still has significant technical and social limitations, some of which are acknowledged in the forecasting literature. Frey and Osborne consider three “engineering bottlenecks” when calculating the automatability of American occupations, identifying “perception and manipulation,” “creative intelligence,” and “social intelligence” as areas that elude technological capability.\textsuperscript{14} Levy identifies broader limitations, arguing that AI will be able to better compete against human labor in tasks that are (a) narrow, such that the data the models use contains most of the contingencies it could face in the future, and (b) structured, such that the machine can easily identify consistent patterns in the data.\textsuperscript{15} Much like the factors described in the ALM model, however, these boundaries are elastic; both future changes in the capabilities of AI-driven automation as well as in the nature of the tasks themselves will continuously shift the window of automatability.

Some forecasts peering through today’s window of automatability nevertheless predict grim outcomes for employment. In their occupation-focused model, Frey and Osborne calculated probabilities of computerization for 702 occupations by using administrative data about the task content of those jobs from the U.S. Department of Labor and having AI experts classify the tasks according to their technical automatability.\textsuperscript{16} The study estimated that 47 percent of U.S. jobs were at high risk (which they defined as a 70 percent chance) of automation within twenty years—and most of these in low-wage occupations. The Frey and Osborne forecast has been extremely influential, dominating the narrative in both the popular press and in subsequent academic work (amassing 3,600+ citations as of the time of this writing).

The More Complicated Reality

Risk calculations like Frey and Osborne’s are often used to predict massive unemployment due to advances in AI. But these forecasts are significantly more complicated than they are sometimes portrayed, in large part due to crucial nuances in how work is executed and how industries are organized. First, and most crucially, technological capability to automate certain tasks does not necessarily translate to the actual automation of those tasks, nor of the occupations that to date have been chiefly comprised of those tasks.
The Future of Work in the Age of AI: Displacement or Risk-Shifting?

These forecasts tend to focus exclusively on technical feasibility, with no account of social, legal, political, or organizational factors.\(^ {17}\) But technologies do not operate in social vacuums, and firms’ adoption and implementation of technologies are contextually dependent on factors like internal organization,\(^ {18}\) institutional and regulatory landscapes,\(^ {19}\) degree of unionization,\(^ {20}\) and other variables.

Importantly, social and political factors have historically affected the distribution of automation risk. In particular, race and ethnicity in the United States can affect whose work is protected from automation and whose is not. For instance, historically, although the artisans whose work was desklined and automated in the first American industrial revolution were largely white, the dangerous, low-wage factory labor that grew as a result of industrialization was largely performed by immigrants and nonwhite workers. Likewise, when considering Frey and Osborne’s predictions in conjunction with racial and ethnic demographic data, it appears likely that white workers are disproportionately more automatable.\(^ {21}\) But white workers continue to have greater social and political leverage along with higher labor market power, thus altering how these demographic groups could be affected by automation.\(^ {22}\) For instance, the predicted polarization of the labor market into low-wage service work and high-wage “knowledge” labor is likely to have different outcomes depending on workers’ race or gender. During this polarization process, black and Hispanic workers competing with white workers for low-wage service work may experience greater job loss due to structural disadvantages like reduced labor market power.\(^ {23}\)

Moreover, automation often leads not to the elimination of occupations, but to changes in their task composition. Using the same framework as Frey and Osborne, but focusing on time spent doing tasks that are capable of automation using current technology, a McKinsey analysis argued that fewer than 5 percent of American jobs can be “entirely” automated.\(^ {24}\) The McKinsey model ultimately makes a convincing argument that AI portends redefinition of human occupations rather than the replacement of entire jobs. This redefinition has occurred repeatedly during previous periods of rapid technological change. ATMs are often cited as an example of the scale effects of new technology outweighing substitution effects of automation: ATMs did not wholly eliminate the need for bank tellers, but rather changed the tasks associated with the role and allowed for the cost-effective expansion of bank branches.\(^ {25}\) As Autor describes in a seminal 2015 work, whether this will be the case in the current wave of AI-driven automation is dependent on a combination of factors like whether nonautomated, “complementary” tasks are easily available elsewhere in the labor market.\(^ {26}\)

Finally, there are limitations to conceptualizing occupations merely as baskets of discrete executable tasks. Though we may distill occupations to their component tasks for purposes of analyzing them, anyone who has held a job knows that work depends on deep-seated human knowledge that cannot always be boiled down to rule-sets and protocols (even nonroutine ones). The anthropologist Michael Polanyi called this the tacit dimension of human knowledge—there are things humans know and do in the course of everyday life that evade easy categorization and can barely be articulated, let alone auto-
These dimensions of human work are hard to capture in economic models, but represent reasons it will be more difficult for machines to wholly assume the roles of human workers. One 2016 OECD analysis applied much of the framework of Frey and Osborne but used self-reported information on the things workers actually do in their given occupation, finding greater variation of tasks within an occupation as well as more groupwork and face-to-face interaction in jobs. This study ultimately estimated that only 9 percent of individuals were at high risk of automation within the next two decades, in contrast to Frey and Osborne’s much more dire forecast.

Another important complication to these forecasts is that they do not attempt to account for indirect forms of worker displacement that might be wrought by AI. These studies focus exclusively on the technical automatability of tasks within particular occupations, but do not account for broader industry-level effects that may more fundamentally restructure labor markets and types of work. A notable example is the booming growth of online retail, supported and enabled by implementation of intelligent supply-chain systems, and the subsequent “retail apocalypse” closing down brick-and-mortar stores across the United States. By one forecast, 75,000 stores are expected to close by 2026, while 25 percent of retail sales are estimated to take place online, up from 16 percent today. Moving retail online does not necessarily directly automate the tasks required from a department store sales associate, but rather eliminates the need for that role altogether, while potentially creating different jobs at other points in the supply chain. The ensuing importance of warehouses over brick-and-mortar stores also creates a space where tasks can be simplified in order to better accommodate the application of AI and robotics. For instance, because it is challenging for robots to safely pick up variable items that have an unpredictable weight or shape—something that comes instinctively to humans—e-retail companies like Amazon are implementing systems that use AI to build appropriately sized boxes around items rather than having a robotic arm pick them up and place them in a box. As Frey and Osborne themselves note, tasks can be changed to become more automatable; indirect unemployment due to AI often results in this task simplification, by taking people out of the equation and instead creating environments more amenable to machines.

Each of these limitations demonstrates a way in which the outcomes of these forecasts are more complicated than they initially appear. It is not clear to what extent AI will displace existing jobs. What is more certain and more imminent is that AI will impact the conditions of work. Rather than focusing on the quantity of displaced work, we ask here how AI might impact the quality of work for workers on the job, by considering how managers leverage intelligent systems to further firms’ objectives. Questions like these are less amenable to broad economic forecasting and breathless headlines—but inarguably, AI’s impact on workers in the here and now has less to do with displacement, and more to do with integration into existing labor structures and managerial practices. Specifically, as we discuss in the next section, AI’s primary effect on work in these contexts is to shift risks previously absorbed by firms onto workers.
AI as Risk Reallocator

Technology has long held the promise of making work more efficient. Technological advances in the workplace are vaunted for their ability to increase productivity, to incentivize "good" work behaviors, to find and eliminate bottlenecks, and the like. By measuring and monitoring and analyzing and predicting, the rhetoric goes, we can find waste, streamline processes, and eliminate superfluous work. The mantra of analytics is practically an article of faith among managers, who believe that data will reveal the secrets to greater profit margins. In this scheme, workers’ labor is an input to be collected, analyzed, and algorithmically optimized like any other. These practices are rooted in the principles of Taylorism, Fordism, and scientific management, each of which aimed to minimize wasted effort and maximize production through the fine-grained pacing and control of work processes. Al in the contemporary workplace follows in the footsteps of this ethos via intensive monitoring and predictive analysis of nearly all aspects of work tasks and the broader supply chain.

Does all this monitoring and analysis make the workplace more efficient? Maybe—but not necessarily because these practices are actually eliminating waste or increasing productivity. Instead, these technologies can insidiously hide work by offloading its burdens from a firm onto its (comparatively less powerful) workers. Lots of inefficiencies still exist in monitored workplaces, but AI-driven managerial practices redistribute the risks and costs of these inefficiencies to workers while serving a firm’s bottom line. We enumerate an illustrative (but nonexclusive) list of four such practices in the following.

Staffing and Scheduling

Traditionally, the risks of fluctuating consumer demand have been borne largely by the firm. Some hours at a store or restaurant, for instance, may be unexpectedly slow. Though managers ideally try to match customer demand to labor supply (i.e., workers on shift), they previously could do so only approximately, usually based on historical indicators like aggregate sales volume during a given period. This often meant that managers bore the risk of overpaying for excess labor capacity (i.e., wages) for unexpectedly slow periods.

Algorithmic technologies have changed the landscape of staffing and scheduling, however, transferring the burden of demand uncertainty from the firm to the worker. More sophisticated staffing algorithms integrate many more sources of data—including, for example, real-time customer traffic derived from in-store sensor networks, as well as external variables like weather—to predict customer demand and associated staffing levels, and to do so more dynamically. The result for workers has been a variety of “just-in-time” scheduling practices that introduce significant precarity and instability into the lives of low-wage workers. These include patterns like irregular and “split-shift” scheduling (i.e., having workers work multiple shorter shifts during periods of high demand, and clocking out in between—leaving that time unpaid); high-fluctuation work schedules (many hours one week, few the next); and short-notice scheduling, including “on-call” shifts (in which
workers must make themselves available for a shift but are notified only just prior to the shift’s beginning about whether they should come in). The effect of each of these practices is to destabilize workers’ livelihoods by interfering with nonwork activities—like school, childcare, or a second job—and creating severe financial stress, leading even to intergenerational cognitive harms. Moreover, these costs are disproportionately borne by women and workers of color, who occupy service positions at higher rates. While firms may lower labor costs due to reduced risk of overstaffing, the upshot of all of these practices is that the burden of the uncertainty of demand is shifted to the workers subject to scheduling systems.

**Defining Compensable Work**

As firms gain more visibility into and control over workers’ activities, they can more narrowly define work to include only very specific tasks and then pay workers for those tasks exclusively. Managerial technology allows firms to focus closely on what is considered essential to a job. The Fair Labor Standards Act (FLSA) requires employers to pay employees for time worked, but only for those activities that are considered “integral and indispensable” to the principal tasks of a job. Under this standard, courts have ruled several activities noncompensable, like commuting to work, waiting to go through required security screenings, and donning and doffing protective gear, even though the principal work tasks cannot, practically speaking, be completed without them. Though many workers (including most gig economy workers) are not covered by the FLSA, the law’s narrow framing of compensable work is conceptually instructive here. Algorithmic technologies may further circumscribe firms’ definitions of essential and compensable work, but they do not actually reduce the amount of work that workers do.

For example: drivers for Uber and other ride-share companies are paid only for the time they are actively transporting a passenger—not the time they spend driving around waiting for the app to alert them to a passenger nearby; not the time they spend driving to a pickup point; not the time they spend returning from a long trip out of town; not the time and expense required to clean their cars and offer amenities in order to get high customer ratings (which can impact the security of their employment). Because these undertakings are not seen as directly generating revenue for the company, they are unpaid. Of course, in reality, all of these tasks are part and parcel of doing the work of Uber driving, and the costs of that work (including both opportunity costs—the time the driver could be making money otherwise, or doing something else entirely—and direct costs, like gas and vehicle wear and tear) are borne entirely by the driver. Though this model of payment isn’t created by algorithmic dispatch—it has, for instance, long been a feature of the truck-driving labor model—the use of AI-driven platforms to support these industries broadens and exacerbates these effects.

Granular measurement capabilities can also be used to more explicitly recalibrate compensation schemes in favor of the firm. In 2015, for instance, Amazon changed how it paid some authors of books available on its Kindle platform. Because Amazon’s technology gave it visibility into exactly how many pages of a book readers actually read, it began
compensating authors on a per-page-read basis, rather than by the number of books downloaded—shifting the risk of a boring book to the author.46 Similarly, music-streaming services like Spotify pay artists on a per-track-streamed basis (where a track is “counted” when a listener plays it for at least thirty seconds), rather than by albums sold or tracks downloaded.47 In theory, compensation models like these reward popularity, and implicitly, quality—but in practice, the model is often blamed for “streambait” homogeneity in cultural production, as risk-averse artists conform to styles most likely to generate revenue under the algorithm.48

Collectively, these trends more tightly circumscribe what is considered compensable work by “counting” certain tasks but not others. And by constricting what is considered compensable work and optimizing narrowly for it, AI-driven systems may increase the proportion of work that is considered residual and unworthy of payment, like producing an (ultimately unpopular) song, driving to a passenger pickup, or replenishing mints to ensure a high rating. Those work activities—what Craig Lambert has termed “shadow work”49—don’t disappear just because they aren’t accounted for. Rather, these systems shift these risks and costs from the employer to the worker, who must internalize the very real labor that doesn’t “count.”50

Detecting and Predicting Loss and Fraud

AI may also be used to redistribute the risk of deliberate damage or loss brought to an enterprise by employees purposively behaving against the firm’s interests. This often involves employees violating the law or the terms of employment—whether by stealing merchandise, embezzling money from company coffers, or sharing a secret recipe—or whistle-blowing to bring to light a firm’s illegal or unethical behavior. The principal-agent problem poses inherent risks to running a business, and employers have historically attempted to lower this risk through myriad low-tech and high-tech means. It is the norm for an employer to call references to determine the supposed character of a potential hire and perform background checks for previous criminal convictions. Employees dealing with sensitive or proprietary information are often required to sign nondisclosure and noncompete agreements. The risks are especially prominent in retail, where the product is directly handled by employees, often without supervision: according to the 2018 National Retail Security Survey, approximately 1.33 percent of retail sales—amounting to about $46.8 billion in costs to U.S. retailers—was lost to inventory “shrink,” with employee theft cited as the second-highest cause of shrink after external shoplifting.51 The costs of shrink make retail a natural adopter of loss-prevention technologies and techniques, from the use of CCTV cameras to the maintenance and creation of an industry-wide hiring blacklist of individuals suspected of theft.52

Employers use AI to continue cracking down on the risk of deliberate damage, often by using technologies that continuously track and analyze worker behavior and activity. Loss prevention firms like Appriss Retail offer services that use AI to model employee behavior and flag unusual behavior that could be fraudulent or harmful to the firm.53 Outside of retail, companies similarly monitor employee activity, especially communications.54
leaked list of phrases from 2008 shows Goldman Sachs flagging emails with lines like “clowns managing the fund,” “report the matter to the sec/nasd/nyse,” or “this won’t happen again” for scrutiny.\(^5^5\) London-based firm StatusToday continuously tracks electronic behavior and flags unusual activity, like an employee accessing files they don’t usually access or copying large numbers of files.\(^5^6\)

Loss and fraud prevention, and the use of AI in its service, may seem to be quite reasonable on the part of the firm; after all, few would condone outright theft, and firms seem justified in protecting their assets, ensuring regulatory compliance, and the like. Our goal is not to pass normative judgment on the propriety or advisability of these aims or practices. Rather, we discuss them here for two reasons related to risk-shifting and worker power. First, though these technologies are explicitly framed as reducing the risk to firms of workers’ deliberate malfeasance, monitoring workers for theft and fraud is often practically inseparable from tracking for productivity or efficiency purposes. The same platform advertised to minimize threats to a firm’s security can be (and often is) also used to ensure employees are maximally productive;\(^5^7\) concerns about fraud may be used as a pretext to justify an entire data collection regime, as has been the case in other contexts (e.g., state benefits provision\(^5^8\)). We discuss productivity monitoring in more detail in the next section.

Second, preventing and detecting loss and fraud have specific implications for risk reallocation between firm and worker. These systems are often predictive, meaning that the harm of malfeasance has not actually happened yet. In other words, rather than mitigating actual loss \textit{ex post}, the employer is looking for potential harm \textit{ex ante}. This is a distinction with an important difference for workers. If systems’ predictive accuracy is poor, or if employers are especially risk-averse—say, in a weak labor market in which they have abundant potential hires—these systems may prevent many workers deemed “risky” from being hired at all. In other words, the risk of future deliberate damage is displaced from firms to potential hires. Employers have long based hiring decisions on heuristics that “mark” workers based on characteristics like race or prior incarceration, often making these workers effectively unhireable and precluding economic opportunity.\(^5^9\) Greater use of predictive systems for loss and fraud prevention may further exacerbate these trends, especially for workers who are already disadvantaged. A further complication arises from the nature of the data in theft prevention databases, which are self-reported and shared among employers, often based merely on suspicion (i.e., without substantiation or subsequent criminal charges) and very likely to be inflected with employers’ own biases. (In fact, concerns about the inaccuracies and lack of due process associated with inclusion in such databases have given rise to lawsuits alleging that their use may violate the Fair Credit Reporting Act.\(^6^0\))

\textbf{Incentivizing and Evaluating Productivity}

Finally, intelligent systems are used to measure, assess, and incentivize workers’ performance in the workplace. Like loss prevention, concern about workers putting forth less than full effort is a feature of principal-agent relations; firms take many steps to incen-
tivize workers to expend more labor\textsuperscript{61} and, conversely, may punish workers for perceived shirking. Though worker surveillance for productivity maximization is nothing new, AI-driven systems may extend the practice into new types of workplaces—for example, workplaces like long-haul trucking, previously shielded by such collection by virtue of its geographic diffusion\textsuperscript{62}—and toward more invasive and fine-grained forms of monitoring.

Amazon, for example, has issued “inactivity reports” for its warehouse workers, detecting when workers temporarily stop moving (even for periods as short as one minute);\textsuperscript{63} it currently holds a patent for a wristband that tracks a worker’s movements and speed, buzzing with haptic feedback to direct the worker to the next item.\textsuperscript{64} Workers in Amazon warehouses have reported grueling pressures, including inadequate breaks for using the bathroom and meeting religious needs, and physical and mental health crises as a result of such strenuous conditions.\textsuperscript{65} Leaked corporate documents show that worker supervision and tracking—up to and including termination of employment for insufficient productivity—is handled by an AI-driven system.\textsuperscript{66} Platform-based firms like Uber also use AI to promote driver productivity, using fleet-wide supply/demand predictions and behavioral-economic “nudges” to tailor incentives toward profit maximization.\textsuperscript{67} In customer-facing service jobs like call centers, AI can be used to monitor not only the speed of work but also alignment with behavioral and affective criteria like tone of voice. In retail settings, workers may be incentivized and evaluated based on automated analysis of their interactions with customers on the floor.\textsuperscript{68}

Productivity incentivization is not \textit{a priori} bad for workers; in commission-based work, for example, it may be advantageous for labor as well as management. But in many contexts, fine-grained monitoring erodes trust, dignity, and any sense of privacy from work, reduces workers’ decisional autonomy,\textsuperscript{69} and opens the door to labor exploitation by driving workers to the limits of their physical and mental capabilities. If working to less than one’s full capacity is considered a form of “time theft,”\textsuperscript{70} similar concerns attach here as they do with respect to loss prevention.

As we have described, intelligent systems in the workplace can be used in the service of several managerial techniques. They may enable firms to dynamically schedule workers, minimizing labor costs while creating substantial instability in workers’ lives. Firms may use AI to narrowly redefine work tasks, concomitantly classifying some practically necessary labor as ancillary and noncompensable. They may use it to predict worker theft and malfeasance, potentially resulting in an underclass of “marked” workers deemed too risky to hire. And they may use it to incentivize productivity by removing all slack from work time, perhaps doing serious damage to workers’ physical and mental health. These dynamics were not created by AI; they have been features of labor/management relations for a long time and will likely remain so for a long time to come. But AI may enable firms to more effectively pursue their existing goals through these practices, therefore offloading burdens and reallocating risks from themselves onto workers.
Displacement, Risk-Shifting, and Policy

Policy recommendations for the future of work commonly focus on mitigating the harms of labor displacement, like unemployment, depressed wages, and increased inequality as a result of labor market polarization. And although AI is often framed as a new frontier for policymaking, proposed solutions often focus on strengthening long-standing social institutions. These recommendations include investing in both K–12 and college education (often with a focus on STEM [science, technology, engineering, and mathematics] fields) and retraining displaced workers to provide them with marketable skills for the new economy; bolstering the social safety net through reforms to unemployment insurance and public benefits programs; and (somewhat more controversially) some support for universal basic income programs that would provide unconditional cash guarantees for all individuals, regardless of circumstance.

These policy proposals stand to benefit millions of Americans whether or not their jobs are displaced by AI and represent sound economic investments in the future of work—whatever it may look like. In addition to proposals like these, however, we should also consider what protections we might provide for workers who retain jobs, in order to temper risk reallocation that intensifies management/worker inequity. For example, a number of states and municipalities have taken steps to curtail worker-unfriendly scheduling practices through fair scheduling laws—sometimes in response to the threat of wage theft lawsuits. These laws do things like require managers to announce schedules further in advance, end “on-call” shifts, and create minimum shift lengths. In so doing, they help to recalibrate employers’ ability to shift costs to workers through algorithmic scheduling.

Other worker protections could similarly reallocate some risks back to firms. One clear avenue would be an end to forced arbitration, which often bars employees from litigating claims against their employers in court; proposed reforms like the Arbitration Fairness Act would prevent employers from being able to enforce arbitration agreements in employment disputes. A second route forward includes reforms to worker classification regimes that characterize many platform-based workers as independent contractors rather than employees, therefore removing some protections due to them under labor law (minimum wage, unionization, etc.); such reforms are currently afoot in some states. More broadly, amendments to the Fair Labor Standards Act could be made to include some workers currently exempt from its protections (for example, long-haul truck drivers) —and in some regulated industries, compensation regimes might be modified to more accurately recognize workers’ time and effort. And we might regulate or ban the use of for-profit “retail justice” databases that blacklist potential employees suspected of theft without due process.

One further note is in order. Organizational sociologists have long examined technological interventions into workplaces and their effects on workplace roles and relationships. A key lesson from this work is that technology has no unified set of effects once deployed in a workplace: it can alter new social dynamics or ossify old ones, depending on the conditions surrounding its deployment—including industry structures, broader economic
forces, workplace culture, and institutional mechanisms for governing relations between labor and management. These studies of previous technologies provide a vital lesson: Contemporary forecasting of AI’s impact on workers, and the ethical issues it is likely to bring to the fore, must include concomitant consideration of specific social, economic, and cultural dynamics in a workplace. Any policies put in place to mitigate negative effects must also take these into account. While this observation is a caveat for forecasters and policymakers, it is also cause for optimism: it suggests that there are many firm-level levers that may mitigate the negative dimensions of workplace AI, and that nothing is set in stone.

Perhaps contrary to our call for workplace-specific action, many of the aforementioned policy proposals we identify—in either the displacement-remediation or risk-reallocation buckets—may seem like they are too general, too basic, or have little to do with artificial intelligence specifically. This is because the issues resulting from integrating AI with work are not wholly new, but are instead the continuation of a long line of labor concerns that have endured and transformed throughout the history of industrialized work. But the specter of AI in the workplace does not necessarily spell doom or dystopia; rather, it elucidates the burdens placed on workers, and may bring new energy to creating policies that protect workers for generations to come—ultimately protecting the quality of work, not just its quantity.

**Acknowledgments**

The authors gratefully acknowledge support from the John D. and Catherine T. MacArthur Foundation, New America, and the Cornell University Center for Social Sciences. We are grateful for helpful comments and insights from Joshua Popp and Matthew Sun.

**Bibliography**


The Future of Work in the Age of AI: Displacement or Risk-Shifting?


Notes:


The Future of Work in the Age of AI: Displacement or Risk-Shifting?


(22) Moradi, “Race.”.


The Future of Work in the Age of AI: Displacement or Risk-Shifting?


(26) Autor, “Why Are There Still So Many Jobs?”.


(32) Brishen Rogers, “Beyond Automation: The Law & Political Economy of Workplace Technological Change,” February 4, 2019, https://papers.ssrn.com/abstract=3327608. Rogers reaches a similar conclusion in his analysis of the law and political economy of workplace automation. Like us, he posits that the threat of automation-induced job loss is “overstated” and that the more pressing issues involve managerial techniques, including worker monitoring and algorithmic scheduling. Rogers also thoughtfully points to the relation of workplace data collection to the “fissuring” of the workplace—that is, firms’ outsourcing of key functions to outside contractors.


(35) We focus here on management of already-hired workers, and bracket from our analysis consideration of AI’s emerging role in hiring processes. The implications of AI for hiring are ably analyzed by Miranda Bogen and Aaron Rieke in “Help Wanted: An Exploration of Hiring Algorithms, Equity, and Bias” (Upturn, Dec. 2018), https://


Levy and Barocas, “Refractive Surveillance.”


Vega v. Gasper, 36 F.3d 417 (5th Cir. 1994).


Llorca v. Collier County Sheriff, 898 F.3d 1319 (11th Cir. 2018).


The Future of Work in the Age of AI: Displacement or Risk-Shifting?


(55) Eamon Javers, “You Won’t Believe What Gets an Email Flagged at Goldman: CNBC Has the List,” CNBC (June 16, 2016), https://www.cnbc.com/2016/06/15/you-wont-believe-what-gets-an-email-flagged-at-goldman-cnbc-has-the-list.html. Though the list cited is from 2008 and was rather low-tech in execution, Goldman Sachs has continued this practice with updated search terms.


The Future of Work in the Age of AI: Displacement or Risk-Shifting?


(68) Levy and Barocas, “Refractive Surveillance.”

The Future of Work in the Age of AI: Displacement or Risk-Shifting?


Pegah Moradi

Pegah Moradi, Department of Information Science, Cornell University

Karen Levy

Karen Levy, Department of Information Science, Cornell University