ANNUAL REVIEWS

Annual Review of Sociology

Automation and Augmentation: Artificial Intelligence, Robots, and Work

Ya-Wen Lei and Rachel Kim

Department of Sociology, Harvard University, Cambridge, Massachusetts, USA; email: yawenlei@fas.harvard.edu

Annu. Rev. Sociol. 2024. 50:19.1–19.22

The *Annual Review of Sociology* is online at soc.annualreviews.org

https://doi.org/10.1146/annurev-soc-090523-050708

Keywords

artificial intelligence, AI, augmentation, automation, countervailing effects, displacement effects, expertise, inequality, materiality, robot, tacit knowledge

Abstract

This article reviews the literature that examines the potential, limitations, and consequences of robots and artificial intelligence (AI) in automation and augmentation across various disciplines. It presents key observations and suggestions from the literature review. Firstly, displacement effects from task automation continue to persist. However, one should not assume an unequivocally increasing efficacy of technology in automation or augmentation, especially given the declining productivity growth in high-income countries and some large emerging economies in recent decades. Jobs less likely to be negatively impacted are those that require diverse tasks, physical dexterity, tacit knowledge, or flexibility, or are protected by professional or trade associations. Despite countervailing effects, without policy intervention, automation and augmentation could widen inequality between social groups, labor and capital, and firms. Secondly, AI's promise in task automation and labor augmentation is mixed. AI tools can cause harm, and dissatisfaction and disengagement often arise from their opaqueness, errors, disregard for critical contexts, lack of tacit knowledge, and lack of domain expertise, as well as their demand for extra labor time and resources. The inadequate autonomy to override AI-based assessments further frustrates users who have to use these AI tools at work. Finally, the article calls for sociological research to specify conditions and mechanisms that ameliorate adverse consequences and enhance labor augmentation by embedding the study of automation and augmentation in concrete social and political contexts at multiple levels.

INTRODUCTION

In recent years, the hype and fear surrounding robots and artificial intelligence (AI) may evoke déjà vu for those familiar with the history of technology and work. Technological advancements since the Industrial Revolution have brought about increased productivity and innovation, along with concerns about job displacement. The debate over automation and its negative impacts has been evident in public discussions and political agendas in the United States and the United Kingdom since the 1960s (Lei 2023). Indeed, our bibliographic search reveals that the term "future of work" has been a prominent theme in book and article titles, as well as academic conferences, since the early 1970s, showing its enduring relevance and importance.

In the recent surge of interest, robotics has become closely associated with the fourth industrial revolution, also known as Industry 4.0. It represents the digital transformation of the industrial sector, integrating advanced technologies like robotics and real-time data analytics to boost productivity and flexibility in manufacturing processes (Pfeiffer 2016). AI has generated significant excitement as a general-purpose technology, comparable to the steam engine and the Internet, with potential to impact various sectors. Computer scientist John McCarthy's (2007, p. 1) widely cited definition describes AI as "the science and engineering of making intelligent machines, especially intelligent computer programs." Machine learning, a subset of AI, and deep learning, a subfield of machine learning, have particularly sparked interest as prediction technologies due to their versatile applications, from natural language processing to recommendation engines, speech, and facial recognition (Dargan et al. 2020). The release of ChatGPT (where GPT stands for generative pre-trained transformer) and other generative AI tools has further fueled the excitement surrounding AI.

The excitement and concerns surrounding robots and AI primarily revolve around automation—the reduction of human input in certain tasks by machines (Agrawal et al. 2019). Recently, augmentation has also gained prominence in academic discourse on robots and AI. Augmentation refers to integrating technology into human workflows to enhance capabilities and productivity. Technology can have both replacing and enabling effects. Scholars emphasize substitution in automation and complementarity in augmentation, respectively (Acemoglu & Restrepo 2020, Brynjolfsson 2022). Though they may appear contradictory, task automation and labor augmentation are not mutually exclusive, as economists Agrawal et al. (2023) point out. In fact, automation in certain tasks can result in labor augmentation in other areas.

Amid the recurring hype and fears surrounding automation-related technology, this article reviews the literature on the potential, limitations, and consequences of robots and AI in automation and augmentation. Scholars studying these topics hail from multiple disciplines, including sociology, economics, human–computer interaction (HCI), law, management, and more. Given the interdisciplinary nature of the subject, sociologists should avoid being insular. Therefore, we have included relevant literature from various fields. The article is divided into three main sections. Firstly, we review the literature that examines the effects of technology on labor demand, wages, employment, and economic and social inequality at the aggregate level. Next, we delve into the literature that investigates the adoption of robots and the implementation of AI technologies for task automation or labor augmentation. Research in this section highlights the importance of studying materiality, tacit knowledge, and expertise to comprehend the potential, limitations, and consequences of robots and AI in automation and augmentation. Lastly, we argue to embed the study of automation and augmentation in social and political contexts, examining how organizational and institutional factors, contexts, and mechanisms at various levels influence diverse outcomes related to the development and utilization of robots and AI.

TECHNOLOGY, LABOR, AND PRODUCTIVITY

The majority of studies on automation examine the impacts of industrial robots on wage and employment because the adoption of AI picked up speed only after 2016 in the United States (Acemoglu 2021). Notably, Acemoglu & Restrepo's (2017) study investigates the impact of increased usage of industrial robots between 1990 and 2007 on local labor markets in the United States. Their findings reveal that areas with higher exposure to industrial robots in manufacturing experienced declines in employment and wages. They also estimate that automation has potentially been the most significant factor in reshaping the wage structure in the United States, accounting for approximately 50% to 70% of the variation in wage changes across demographic groups between 1980 and 2019. Research in the context of the European Union finds that the presence of one additional robot per thousand workers reduces the employment rate by 0.16–0.20 percentage points. The negative effect is particularly evident for workers with middle-level education (Chiacchio et al. 2018). Autor & Salomons's (2018) study of Organisation for Economic Co-operation and Development countries from 2007 to 2011 found that automation displaces employment and reduces labor's share of value-added in the industries in which it originates.

In their efforts to estimate the impact of automation, economists have developed comprehensive frameworks to gain a better understanding of the effects of technology on work and employment. These frameworks encompass the analysis of displacement effects and countervailing effects.

Displacement Effects

Initially, labor economists placed significant emphasis on skill as the primary factor in explaining how technology replaces human labor. However, there has been a notable shift in focus from skill to task among many economists. The skill-biased technical change (SBTC) theory posits that technology favors high-skilled workers and disadvantages low-skilled workers. Since high-skilled workers are capable of using technology to increase production, technological advancement creates demand for them; in contrast, low-skilled workers will be replaced by technology (Katz & Murphy 1992). The SBTC theory corresponds with the upskilling theory in sociology, according to which automation leads to an increase in workers' skills as workers are expected to learn new technologies (Adler 1992, Attewell 1992). However, in contrast to the prediction made by the SBTC theory, labor economists have found that computers and robots mostly displaced workers with middle-skilled cognitive and manual jobs (Autor 2015). Some labor economists further argue that the SBTC theory cannot explain the continued existence of low-skilled jobs and occupational polarization—employment growth in both high-skilled and low-skilled occupations—in economically developed countries (Goos & Manning 2007). These critiques highlight the limitations of the SBTC theory.

Moreover, both the SBTC theory and some of its critiques fail to acknowledge the challenges involved in measuring and comparing skill levels. Sociologists of work have extensively criticized the ambiguities and complexities associated with the concept of skill. These criticisms primarily revolve around questions such as whether skill should be treated as a measurable attribute of individuals or jobs, and the comparability of different skills (Attewell 1990, Spenner 1990). However, we observe that numerous studies across various disciplines continue to differentiate between distinct skill levels and persist in utilizing terms such as low-skilled workers and high-skilled workers.

The limitations of the SBTC theory have led labor economists to propose task-based approaches that highlight task as the central unit of production and consider both displacement and countervailing effects of technology (Acemoglu & Restrepo 2019a). When examining displacement effects, labor economists consider the comparative advantages and relative productivity of labor and capital in different tasks. They posit that when capital is sufficiently inexpensive or productive at the margin in certain tasks, automation will substitute capital for labor in those tasks previously performed by human labor, leading to the displacement of workers. While increasing output per worker, such displacement effects could lead to a decrease in labor demand, wages, employment, and the share of labor in national income (Acemoglu & Restrepo 2019a).

Among labor economists who take a task-based approach, some have attempted to specify the types of tasks that are more likely to be automated by technology. A group of economists assert that technology can replace human labor in routine tasks—activities that can be expressed in step-by-step procedures—but not nonroutine tasks (Autor et al. 2003, Goos & Manning 2007). Routine tasks are characterized by middle-skilled cognitive and manual activities, such as many tasks that were once performed by skilled artisans in factories and office clerks. Nonroutine tasks encompass high-skilled cognitive and low-skilled manual activities. Nonroutine manual tasks, such as cleaning, make up low-skilled jobs in the service sector. Labor economists further argue that the material constraint of technology in substituting for nonroutine tasks and the movement of people from routine, middle-skilled jobs to nonroutine, low-skilled jobs can lead to occupational polarization and wage inequality (Autor et al. 2003, Goos & Manning 2007, Van Reenen 2011).

In the era of AI, AI technologies have the capacity to automate various cognitive tasks typically associated with intermediate-level skills in white-collar professions. Some scholars go even further to argue that AI technologies can routinize high-skilled work, potentially resulting in the decline of professionals in fields such as accountancy, law, health, and architecture (Susskind & Susskind 2015). We delve into empirical research regarding this aspect later in the article.

It is also important to highlight that while nonroutine, low-skilled jobs are generally less susceptible to displacement by technology, the advancement of technology can still result in an increase in nonstandard forms of employment, underemployment, and a decline in job quality (Acemoglu 2021). For example, information and communications technologies and AI enable the automation of matching service providers with consumers and managing nonroutine, low-skilled tasks. In the platform economy, a significant number of workers function as independent contractors, confronting increased technological surveillance and control along with insufficient social protection (Griesbach et al. 2019, Lei 2021, Vallas 2017). Research on occupational changes in the United Kingdom also reveals that although the evidence does not suggest workers there currently face an immediate risk of technological unemployment, technology does diminish the quality of jobs (Spencer & Slater 2020).

Countervailing Effects

In addition to displacement effects, economists have emphasized the countervailing effects of technology, which can expand labor demand. Empirical studies support the existence of countervailing effects. For instance, while some research demonstrates negative impacts of robotization on overall employment (Acemoglu & Restrepo 2017), other research finds positive effects (Autor & Salomons 2018). A sociological study investigating the influence of robotization on various occupational types in the United States reveals that increased industrial robot usage is associated with growth in high-skilled, nonroutine jobs as well as middle-skilled, routine, and manual jobs (Dahlin 2019).

According to labor economists, a primary source of countervailing effects is the productivity effect. This effect arises from automation technology reducing the cost of performing certain tasks, leading to increased demand for labor in nonautomated tasks across sectors that are undergoing or not experiencing automation (Acemoglu & Restrepo 2018, Autor 2015). The second source of countervailing effects originates from capital accumulation. Acemoglu & Restrepo (2019a) argue that the significant demand for capital in the automation process can lead to further capital accumulation (for example, by raising the rental rate of capital), thus bolstering the demand for labor. Countervailing effects can also emerge from the deepening of automation, which involves technological improvements in tasks that have already been replaced by capital. This type of automation can boost the productivity of capital, leading to an increase in labor demand without generating additional displacement effects. However, the countervailing effects might not be strong enough to establish a balanced growth path, which implies that technology could still lead to more of an increase in output per worker than in wages and a decrease in the share of labor in national income (Acemoglu & Restrepo 2019a).

Drawing on historical examples, labor economists argue that the emergence of new labor-intensive tasks, where labor holds a comparative advantage over capital, is the most significant factor in balancing the growth process and mitigating displacement effects. They refer to this countervailing effect as a reinstatement effect. As new labor-intensive tasks reintegrate labor into the production process, they contribute to increased employment and/or wages (Acemoglu 2021). Acemoglu & Restrepo (2019b) highlight a notable increase in wages after World War II due to rapid automation in specific tasks and the introduction of numerous new tasks, countering the potential negative effects of automation on the labor market.

Indeed, the adoption of robots and AI has led to the emergence of new tasks and occupations. In China, for instance, the government has classified and described various emerging occupations, including industrial robot system operators, industrial robot system maintenance technicians, service robot application technicians, AI engineering technicians, big data engineering technicians, and AI trainers (SIPAC 2021). Research indicates that with the integration of AI, newly arising job roles involve training and interacting with the AI system, communicating its functionalities to customers, and monitoring and sustaining its performance (Acemoglu & Restrepo 2018, 2019a).

Emerging tasks have varying compensation levels. Low-paid tasks, exemplified by "ghost work" identified by Gray & Suri (2019), involve workers in poorly paid and often overlooked roles. These workers address glitches in automated tools and ensure the smooth functioning of mobile apps, websites, and AI applications. Data labeling work, considered labor intensive and time consuming, requires minimal training and is often performed by workers in economically developing countries (Kshetri 2021, Shestakofsky 2017). On the other hand, high-paying occupations in the AI field involve tasks like designing and applying AI algorithms, testing and fine-tuning them, and developing AI solutions. These roles require analytical, adaptive, interpersonal, and communication skills (Trajtenberg 2019). Such well-paid professional occupations are typically found in economically developed countries or major cities within developing countries. For a comprehensive understanding of the occupational stratification in the era of algorithms, Burrell & Fourcade (2021) offer an extensive review article.

The emergence of new tasks does not ensure a smooth labor market adjustment. Displacement effects from automation can profoundly impact workers employed in automated tasks, leading to challenges in their adaptation. Simultaneously, there might be a shortage of workers with the required skills to perform newly arising tasks in response to changing technological conditions (Acemoglu & Restrepo 2019a).

Implications for Economic and Social Inequality

The literature also indicates a correlation between automation technology and the widening inequality in various aspects, particularly inequality between different social groups, between capital and labor, and among different firms.

Inequality between social groups. The effects of automation technology, including displacement effects, countervailing effects, underemployment, and declining job quality, can disproportionately impact individuals based on race/ethnicity, gender, education level, industry, and region. Vulnerable workers in easily automated occupations are more likely to be adversely affected.

Generally, individuals with lower levels of education are more vulnerable to these adverse impacts (Katz & Murphy 1992). Acemoglu & Restrepo (2017) state that men's jobs, particularly those with lower educational requirements and involving routine industrial tasks, are particularly susceptible to automation. Yet, in service industries and the public sector, a different pattern emerges. For example, the UK Office for National Statistics notes that automation in these sectors, such as self-checkouts at supermarkets and paralegals in law firms, tends to affect women and young workers the most. Additionally, the likelihood of job losses varies by locality, with peripheral regions outside main growth centers experiencing greater susceptibility to job displacement due to automation (Clifton et al. 2020).

In contrast, countervailing effects, particularly reinstatement effects, can alleviate adverse consequences of automation by creating new tasks and job opportunities. However, these emerging roles often require higher education and skills, benefiting individuals with such qualifications (Acemoglu & Restrepo 2018, 2019a). Furthermore, research shows that AI development and adoption vary by region, closely linked to the strengths of tech sectors in different areas. Industries and occupations tend to cluster geographically, leading to spatial inequality at various levels, including regional disparities within countries and inequalities across nations (Clifton et al. 2020, Raj & Seamans 2019).

Inequality between capital and labor. The adoption of technology, particularly automation, can widen inequality between capital and labor. Scholars have observed a decline in labor's share of national income and increasing returns to capital owners, starting in the 1980s and 1990s in advanced and developing economies, respectively, and becoming more pronounced in the 2000s (Dao et al. 2017a). Some economists link the impact of technology to the erosion of labor's share of national income, especially in countries and sectors more specialized in routine-intensive activities (Dao et al. 2017b). Autor & Salomons's (2018) analysis of US data from 1970 to 2007 found that automation technologies, as embodied in total factor productivity growth, have been employment-augmenting yet labor share–displacing over the past four decades. While the countervailing effects of technology were stronger in terms of overall employment, they did not fully offset the displacement effects concerning labor share changes. In summary, the adoption of technology by capital owners has led to workers receiving a diminishing share of the economic pie.

Inequality among firms. Despite the rising returns for capital owners, not all firms have equal opportunities to develop or adopt automation-related technology (Acemoglu & Johnson 2023). In China, larger electronics manufacturers are more likely to invest in robot installations compared with smaller counterparts, primarily due to the capital requirements involved (Cheng et al. 2019). In the realm of AI, data serve as the essential fuel for prediction-driven tasks. Substantial capital investments are necessary to capture, compile, and process extensive volumes of data. Consequently, firms with advantages in these areas, such as those possessing extensive digital infrastructure or platforms, are more inclined to develop and employ AI technologies (Ciuriak 2018).

Research suggests that the growing dominance of large firms, including those in the tech sector, can lead to widening inequality among firms in their ability to innovate and employ AI technology. The rise of superstar firms, known for superior productivity, is attributed to globalization and technological advancements, leading to increased sales by these firms across various industries. This domination of industries by superstar firms results in high markups and a lower labor share of value added (Autor et al. 2020). Notable examples of such firms in the tech sector include Google, Microsoft, Apple, Meta, and Amazon. Superstar firms have a significant impact on the decline of labor's share of national income and leave limited resources for their competitors, raising concerns about antitrust issues (Balliester & Elsheikhi 2018, Ciuriak 2018). Moreover, conglomerates with significant economic resources and technological prowess can leverage their advantages into political influence, shaping institutional and policy frameworks to further their own interests, potentially neglecting the welfare of the general public, labor, consumers, and other enterprises (Acemoglu 2021). This trend is evident in countries like the United States and China (Lei 2023, Zuboff 2019).

The Productivity Paradox

Despite extensive literature on technology's impact on labor demand and employment, social scientists face a paradox. On one hand, technology, particularly automation technologies, has the potential to enhance productivity. On the other hand, actual productivity growth measured statistically does not always align with these expectations. For instance, labor economists have found low aggregate productivity growth in the United States since 1974, except for the period between 1995 and 2004 when it rebounded due to the widespread adoption of information technology (IT) (Gordon 2016). This trend of slowing productivity growth is not limited to the United States but is also observed in other high-income countries and large emerging economies like China (Syverson 2017). In spite of China's efforts to promote a science- and technology-driven economy and extensive adoption of automation technologies (Lei 2022, 2023), the country has experienced a marked slowdown in productivity growth since the 2008 global financial crisis (Brandt et al. 2020).

The productivity paradox has sparked debates among scholars. Some argue that it results from output productivity mismeasurement, as technology's benefits are not fully reflected in aggregate statistics (Mokyr 2014). However, others find this hypothesis unlikely, given the pervasive nature of the productivity slowdown (Acemoglu & Restrepo 2019a, Brynjolfsson et al. 2019). Additionally, the paradox is further complicated by the fact that the slowdown is more severe in IT-intensive industries than in the rest of manufacturing (Acemoglu et al. 2014).

Scholars have proposed two alternative explanations. The first focuses on the concept of excessive automation, which refers to a faster rate of automation than what is socially desirable. Excessive automation can create inefficiencies by wasting resources and displacing labor. Government subsidies that favor capital over labor (e.g., in the form of tax codes) and labor market imperfections can contribute to an inefficient increase in automation, leading to the misallocation of capital and labor (Acemoglu & Restrepo 2018). Acemoglu & Restrepo (2019a) also introduce the concept of so-so technologies, which pertains to technologies that are sufficient for adoption but lack the potential to significantly enhance productivity. The adoption of such technologies can result in limited productivity effects, excessive automation, and adverse social consequences. Sociological research offers evidence that supports this line of argument. For instance, in China, certain manufacturers, driven by government subsidies and additional incentives, failed to adequately consider factors such as productivity growth, actual outcomes of automation, and social consequences when installing robots (Lei 2022).

The second explanation is underinvestment in non-automation-related technologies and inadequate investment in complements to automation technology. Acemoglu & Restrepo (2019a) argue that the heavy focus on AI, machine learning, and big data techniques for task automation could lead to reduced investment in other productivity-enhancing technologies. Brynjolfsson et al. (2019) highlight the time lag between the emergence of a technology and its measured productivity effects. They emphasize the challenges and adjustment costs at individual, organizational, and societal levels in developing complements for automation technology to fully realize its benefits. These complements include not just human capital and skills but also novel processes, business models, organizational structures, and cultural changes.

TECHNOLOGY AT WORK

Parts of the previous discussions, particularly those addressing the productivity paradox and the uncertainty surrounding countervailing effects, underscore the significance of analyzing the development and implementation of AI, robotics, and associated technologies in the workplace. Amid the ongoing hype and fear surrounding robots and AI, sociologists and other social scientists have focused on scrutinizing their adoption at work, revealing dangers, pitfalls, limitations, and potential.

Dangers and Pitfalls: Surveillance, Control, Manipulation, and Bias

While the economics literature highlights the adverse effects of automation-related technology on wages and employment, research in sociology and other disciplines has also delved into the additional detrimental consequences of AI technologies for workers, marginalized groups, and citizens as they are developed and implemented in the workplace. These consequences can arise either unintentionally or intentionally.

As mentioned in the preceding section, a growing number of firms use AI technologies to automate managerial surveillance and exert greater control over workers (Brayne & Christin 2020, Griesbach et al. 2019, Kellogg et al. 2020, Rosenblat 2018, Vallas 2017). Among workers, those in low-skilled positions often experience the most intensive and extensive technological control since work supervision can be automated relatively easily (Lei 2023). However, this heightened control has resulted in various forms of resistance from workers, as Kellogg et al. (2020) systematically theorized. Research has also shown individual efforts and collective action organized by workers to resist automated managerial surveillance and control (Brayne & Christin 2020, Lei 2021, Tassinari & Maccarrone 2020).

Furthermore, AI technologies have evolved into not only instruments utilized by employers but also political actors. Authoritarian governments, such as the Chinese government, have adopted AI technologies to surveil their citizens, particularly targeting ethnic minorities, and to carry out censorship and propaganda (Chin & Lin 2022, Munn 2022a). There are also concerns among scholars about AI technologies encroaching upon privacy and individual autonomy, as well as being exploited by political actors to influence the public through manipulating information, spreading misinformation, and disseminating disinformation. As a result, scholars argue that AI poses a significant threat to democracy (Acemoglu 2021, Zuboff 2019).

Another major concern associated with the use of AI technologies at work is their potential to perpetuate bias and social inequality (Acemoglu 2021; Benjamin 2019; Broussard 2018, 2023; Eubanks 2018; Ferrer et al. 2021; Joyce et al. 2021; Noble 2018). AI tools, designed to enhance productivity and efficiency, are being increasingly utilized to automate and assist human decision-making across various domains. These include policing, the judicial system, finance, business, government and social services, healthcare, education, and more. While AI-assisted decision-making may appear neutral, objective, and superior to human decision-making in terms of efficiency and susceptibility to bias, research has extensively shown how it can inadvertently lead to prejudice or bias against individuals or groups based on their inherent or acquired characteristics (Mehrabi et al. 2021). When applied to determine the allocation of life chances, including opportunities, resources, and restrictions, AI-assisted decision-making can perpetuate and exacerbate social inequality. To exemplify this issue, we provide two instances.

Research shows that facial recognition technology exhibits higher accuracy rates when identifying faces with lighter skin and traditionally male facial features, compared with darker skin and traditionally female facial features (Buolamwini & Gebru 2018). The increasing use of this technology in law enforcement can lead to disproportionate misidentifications, raising concerns about civil liberties for women and communities of color (Schuetz 2021). Scholars have also discovered that a risk-assessment algorithm used in healthcare systems tends to assign the same predicted risk score to healthier White patients and less healthy Black patients, despite their different health conditions. This occurs because the algorithm uses healthcare expenditure as a proxy for health needs, resulting in disparities in resource allocation and healthcare provision, thereby reinforcing existing inequalities (Obermeyer et al. 2019). These examples are just a few of many studies highlighting similar insights.

Existing studies identify multiple sources of bias in AI-assisted decision-making. The first arises from biased data generation used in training and analysis, where historical discrimination and unrepresentative data sets can introduce biases into algorithms. The second source of bias lies in algorithm design choices, which can either introduce new biases or amplify existing ones in the data. Additionally, during AI system implementation, biased algorithm predictions can influence user decisions, creating a feedback loop that generates more biased data for training future algorithms (Gerdon et al. 2022, Mehrabi et al. 2021).

Various actors, including scholars, journalists, and activists, have responded to the dangers and drawbacks of AI technology with criticisms and actions. In both the public and private sectors, there has been a proliferation of published AI guidelines and codes of ethics (Greene et al. 2019, Munn 2022b). Tech companies have hired personnel to institutionalize and enforce ethical guidelines (Metcalf & Moss 2019). Furthermore, there has been a rise in activism and scholarly initiatives aimed at promoting fairness, transparency, and accountability in AI technology.

However, the effectiveness of these efforts remains uncertain and disputed. Research highlights the disparity between ethics as a mode of normative inquiry and ethics as a practical undertaking. The practical application of tech ethics encounters several significant constraints. Tech ethics is marked by vagueness and ineffectiveness (Munn 2022b). While its implementation primarily centers on AI practitioners (e.g., data scientists and engineers) and the design of technology, AI practitioners encounter difficulties in selecting performance metrics, identifying the most pertinent direct stakeholders and demographic groups to focus on, and gathering data sets for evaluations. There is insufficient interaction between AI practitioners and direct stakeholders or domain experts. Moreover, the implementation of tech ethics becomes assimilated into the motives and rewards of corporate frameworks that prioritize customers over marginalized groups (Madaio et al. 2022). These limitations suggest that technology ethics can lead to a phenomenon known as ethics-washing within corporations. This involves adopting ethical terminology to divert scrutiny and avoid governmental regulations, all the while sidestepping a genuine dedication to ethical conduct (Attard-Frost et al. 2023, Green 2021, Metcalf & Moss 2019).

Limitations and Potential: Materiality, Tacit Knowledge, and Expertise

In addition to uncovering the dangers and pitfalls of AI, sociologists and other social scientists also explore the limitations and potential of robotics and AI in automation and augmentation. Due to their distinct material properties, we discuss research on these two technologies separately. However, as we illustrate later, materiality (i.e., the physical attributes of technology and human body), tacit knowledge, and expertise are consistent and interrelated themes in the literature.

Robots. Research on robotic automation suggests that terms such as “routine” and “low-skilled” underestimate the critical significance of living human labor capacity and the contributions of tacit knowledge (Lei 2022, Pfeiffer 2016). The concept of living human capacity focuses on the somatic affordances of the human body in comparison with the capabilities of automation technology. Such affordances enable humans to understand certain processes without explicit knowledge of how those processes work. Scholars have long emphasized the role of tacit knowledge in society. Explicit and tacit knowledge differ in terms of codifiability, mechanisms for knowledge transfer, and methods of acquisition and accumulation. Tacit knowledge is often intuitive, unarticulated, and action oriented, making it challenging to express through communication. The transfer of tacit knowledge often requires close interaction, shared understanding, and practical experience in the relevant context (Lam 2000; Polanyi 1958, 1967).

Collins (2010) classifies tacit knowledge into three categories based on the challenges in codifying it. Weak, or relational, tacit knowledge remains concealed due to organizational boundaries within societies. However, with sufficient effort, relational tacit knowledge can be made explicit. Medium, or somatic, tacit knowledge is embodied in the human physique but not fundamentally inexplicable. Strong, or collective, tacit knowledge is deeply embedded in the socially shared world of meanings, and Collins argues that it cannot be fully made explicit or integrated. Even when tacit knowledge can be made explicit, it may not be economically viable (Shestakofsky 2017). Understanding tacit knowledge helps analyze why and how a machine functions or falls short in specific contexts (Bijker 2011).

Scholars have examined classic examples of routine work against the backdrop of Industry 4.0. Major global markets for industrial robots encompass China, Germany, Japan, the United States, and South Korea, with substantial revenues generated in the automotive and electronics manufacturing industries (Cheng et al. 2019). Consequently, a significant portion of research on robotic automation has been conducted within these contexts. Despite advancements, studies of electronics and automotive manufacturing show certain actions remain challenging for robots to replicate, particularly those relying on sensory perception and body memory. These actions involve somatic affordances of the human body, making it difficult to transfer somatic tacit knowledge in a cost-efficient manner (Lei 2022, Pfeiffer 2016). For instance, knowing the appropriate amount of force and pressure to apply through the sense of touch poses challenges. In light of these challenges, major electronics manufacturers in China have once again recognized the unique qualities of human workers in precise, skilled, and flexible assembly tasks, even though such tasks are usually perceived as routine or low-skilled (Lei 2022).

Research also suggests that the prevailing scholarly and popular discourse, which solely emphasizes the perils of robots replacing humans, often fails to acknowledge the emergence of new tasks that demand an increased level of physical dexterity. In her study of electronics manufacturing, Lei (2022) found that the limitations of robots are further magnified by the increasing delicacy of electronics products, which demand higher levels of flexibility and precision during assembly. As a result, there is an adjustment in the division of labor between human workers and machines. Robots are utilized to relieve workers from certain tasks, particularly those that are dirty, dull, or dangerous. This enables workers to focus on assignments that require greater levels of physical dexterity. It is thus not surprising that globally mobile capital persists in seeking cheap and disciplined labor on a global scale.

Furthermore, collective tacit knowledge plays a crucial role in the uninterrupted functioning and continuous improvement of assembly lines. It is not only engineers and technicians who possess such collective tacit knowledge, but also low-skilled workers. As these workers gain experience in performing routine tasks, their acquired knowledge enables them to proactively address issues and adapt to unpredictable circumstances. The increasing complexity of assembly lines due to automation technology amplifies the significance of tacit knowledge in their operation. While human supervision, adjustment, maintenance, and improvement remain essential, workers handling routine tasks must be trained to work alongside the more intricate assembly lines (Bainbridge 1983, Pfeiffer 2016).

Toyota's decision to replace robots with humans amid the growing hype surrounding automation exemplifies the importance of collective tacit knowledge in manufacturing. In a factory in Japan, the company transitioned from automated equipment to manual forging of crankshafts using human workers, resulting in surprising outcomes: a 10% reduction in material waste and a 96% reduction in the production line length. Instead of viewing deskilling through automation as a victory of firms over workers (Noble 1984), the Toyota leadership believed that an excessive reliance on automation could undermine tacit knowledge and lead to a lack of understanding among workers regarding the production processes. This, in turn, could result in a deficiency of continuous improvement in manufacturing, potentially leading to losses—rather than gains—for the company in the long term (Bauer et al. 2018). Toyota's decision corresponds to what Bainbridge (1983) wrote about the ironies of automation. Although routine tasks are the easiest to automate among all kinds of tasks, performing mundane tasks helps workers hone their skills and knowledge, preparing them for more complex tasks. As such, automating routine tasks can deprive workers of the knowledge and experiences they need to perform tasks that machines (e.g., robots and AI) cannot (De Bruyn et al. 2020).

Research has also revealed the limitations of using robots to automate certain tasks and thereby augment human capabilities. The limitations are primarily due to the addition of tasks, reconfiguration of work practices, and opposition from various actors that arise upon the introduction of robots, as illustrated by the incorporation of robots into long-term care facilities for routine tasks (Vogt & König 2021). Japan's rapidly aging society requires addressing the long-term care labor shortage. The introduction of robotic devices has been disappointing and has been accompanied by costly work practice reconfigurations. Even if successfully implemented, robotic devices, like robotic exoskeletons, often require additional adjustments and do not seamlessly integrate into intended tasks. Dealing with these challenges consumes valuable time that could be devoted to patient care. While the adopted so-so technologies may replace human labor for certain tasks, they do not significantly enhance productivity (Acemoglu & Restrepo 2019a). Furthermore, the introduction of humanoid robots to provide services encountered opposition from visitors who felt a sense of alienation in their presence. Research indicates that long-term care facilities in Japan have taken steps to eliminate impractical robotic devices and reconsider their purchasing strategies (Vogt & König 2021).

Artificial intelligence. In general, research suggests AI excels in tasks that have clearly defined parameters, such as prediction, routine decision-making, logistics, and pattern recognition. However, despite their remarkable advancements, AI technologies still face challenges in automating tasks that involve complex reasoning; judgment; abstract problem-solving; and a combination of physical activity, empathy, and communication skills due to their material constraints (Acemoglu & Restrepo 2019a).

Relatedly, challenges in transferring tacit knowledge between AI models and organizations also significantly limit the capacity and application of AI in task automation or labor augmentation. On one hand, the absence of common sense and collective tacit knowledge makes the specification of a holistic objective function and AI-powered decision-making particularly complex (De Bruyn et al. 2020, Shestakofsky 2017). For instance, social workers employ AI-assisted tools to evaluate the risk of child maltreatment. However, AI models cannot adequately capture the rich knowledge and contextual information possessed by social workers, leading to unrealistic assessments of risk (Kawakami et al. 2022). On the other hand, developers and users continue to encounter challenges in explaining, interpreting, and controlling the operation of complex and opaque AI models, which are often perceived as black boxes (De Bruyn et al. 2020). Such difficulties in transferring tacit knowledge tend to be more complex than in the case of robotics, as there are fewer difficulties in explainability and interpretability in the latter.

Amid the hype surrounding AI, scholars have delved into comparing the accuracy and fairness of AI-based automated decision-making with that of human decision-making in areas such as healthcare and criminal justice (Jussupow et al. 2022). Much of this research centers around risk-assessment tools employed within the criminal justice system in the United States. While some studies suggest that algorithms outperform human decision-makers in estimating risk within the criminal justice system (Green & Chen 2019), others indicate that a widely used AI-based risk assessment tool is no more accurate or fair than predictions made by individuals lacking criminal justice expertise. Furthermore, the tool fares no better than a simple logistic regression model when it comes to decision-making (Dressel & Farid 2018). Given the mixed findings in existing literature, scholars suggest that we should evaluate the effectiveness of AI-based and human decision-making on a case-by-case basis (Gerdon et al. 2022).

Also, solely focusing on the accuracy and fairness of AI models can lead to overlooking how the introduction of AI-based tools can reconfigure work conditions and result in additional labor time and expenditures. This is similar to the case of introducing robots to Japan's long-term care facilities (Vogt & König 2021). While many AI technologies aim to reduce the time required for certain tasks and increase productivity and efficiency, research on AI applications in healthcare, public service, and farm management reveals that the implementation and utilization of AI tools often require additional, often invisible and undervalued, labor and expenditures. This encompasses managing the opacity and errors of AI tools, addressing disruptions, performing emotional labor toward other coworkers or clients, and adapting to evolving work routines, logics, and infrastructures. The need for extra labor often adds stress to users who already face time pressure and labor shortages, leading to doubts about the overall benefits of AI tools and concerns about so-so technologies (Acemoglu & Restrepo 2018, Jussupow et al. 2022, Kawakami et al. 2022, Mateescu & Elish 2019). In the case of healthcare, scholars also show how the introduction of algorithmic tools led to a disjuncture between human-initiated care work and work that supports algorithms, making hospitals algorithmically centered rather than human centered (Bailey et al. 2020).

A significant theme that requires further research pertains to professionals, experts, and expertise. While the challenge to and skepticism toward professionals is not a new phenomenon (Epstein 1995), some scholars argue that with the wider application of AI, professionals in fields such as law, accounting, and healthcare are increasingly vulnerable to highly capable machines that can automate complex tasks and redefine the distribution of expertise. As a result, some speculate that AI technologies will render professionals obsolete (Susskind & Susskind 2015). However, contrary to Susskind & Susskind's (2015) sweeping prediction about the obsolescence of professionals, existing studies suggest that AI technologies, while capable of substituting for skilled labor in certain tasks, primarily serve to augment the capabilities of professionals (Clifton et al. 2020, Rodgers et al. 2023).

Among all professionals, radiologists arguably face one of the highest risks of automation. Traditionally, radiologists rely on visual assessment of medical images, but AI technologies now automatically recognize complex patterns and provide quantitative assessments of radiographic characteristics. Presently, AI development has reached a diagnostic performance on par with that of medical experts and has even surpassed clinicians with less experience, particularly in image recognition–related fields (Shen et al. 2019). Despite these advancements, radiologists are experiencing not technological unemployment but rather a global labor shortage due to an aging population and increased healthcare accessibility. AI has emerged as a solution to address this shortage (Chen et al. 2021, RSNA 2022). Currently, radiologists retain the authority to decide whether and how to incorporate AI's diagnostic assessment into their final diagnosis (Jussupow et al. 2022, Lebovitz et al. 2022).

In the realm of legal practice, the integration of AI within law firms is still in its early stages. Empirical research regarding how lawyers and paralegals use AI in their practice remains limited. Many lawyers, including those who have experience using AI, often express skepticism about the efficiency and effectiveness of legal AI technologies. Currently, AI is primarily utilized in legal research to expedite certain tasks, such as identifying potentially relevant information, thereby reducing the time spent on these activities (Brooks et al. 2020). Machine learning algorithms, in particular, aid in the identification of pertinent cases or statutes and generate legal research memos. AI technologies also assist lawyers and paralegals in uncovering relevant information for tasks involving due diligence and contract analytics. Overall, despite technological advancements, legal AI is predominantly utilized in law firms for peripheral tasks in the legal process. This is due to the complex nature of legal reasoning and judgment, as well as the need for communication with multiple parties involved in legal proceedings (Rodgers et al. 2023). The substitution of labor in the information search and administrative process potentially enables lawyers and paralegals to focus on core legal responsibilities within their practice, which require skills such as persuasion, judgment, creativity, and interpersonal communication.

While the adoption of AI technology has not rendered professionals obsolete, it has ushered in new areas and forms of expertise within professional services, particularly services centered around AI. Drawing on the work of Eyal & Pok (2011), we conceptualize expertise as a network comprising actors, instruments, statements, and institutional arrangements. Research indicates a growing number of data scientists assuming roles as technical experts within diverse firms (Dorschel 2021). This trend parallels a larger pattern that encompasses the integration of computer engineers into conventional professional work settings and domains of expertise (Ensmenger 2010, Pardo-Guerra 2019). Nevertheless, data scientists undertake conflicting roles, functioning both as generalists and specialists, technicians and communicators, as well as data exploiters and ethicists. These roles distinguish them from statisticians or computer scientists (Dorschel 2021). For instance, in the retail industry, data scientists are tasked with gathering and explicating tacit knowledge from professionals with domain expertise. They proactively encourage domain experts in their company to explicitly articulate the underlying theories behind their decisions (Valentine & Hinds 2022). Further exploration of the relationship and interaction between emerging technical experts in the field of AI and other domain experts is warranted.

Scholars have observed the rise of technical experts in conventional professional services, who are responsible for tasks like system selection, AI procurement, data curation, and model evaluation (Kluttz & Mulligan 2019, Rodgers et al. 2023). Some law firms in the United States and United Kingdom seek data analysts proficient in data science, Python, SQL, and relevant programming languages for emerging tasks (Sako et al. 2022). This trend raises the possibility of blurred roles and boundaries between various types of experts, such as paralegals, legal assistants, and technical experts. Semiprofessionals, like radiographers, express more concern about AI's impact on their roles and skills compared with radiologists. Conventional professionals, certified and affiliated with professional associations, show greater confidence in their expertise despite AI's rise (Chen et al. 2021). To address the opportunities and challenges presented by AI technologies, professional associations have initiated the development of policies and initiatives within their respective fields. The influence of professional associations can play a substantial role in shaping the configuration of expertise networks (Rodgers et al. 2023).

The emergence of new forms of expertise centered on AI further raises critical questions about the relationship between human expertise and AI-powered expertise. Inquiries in this regard intersect with studies of HCI. In high-stakes decisions, conventional professionals like physicians, judges, and lawyers usually have the autonomy to follow or disregard the assessments made by AI technologies (Jussupow et al. 2021, Stevenson & Doleac 2022). A challenge arises as AI systems often provide limited information for professionals to understand their assessments. Despite this limitation, conventional professionals exercise human oversight and discretion when making ultimate judgments, recommendations, or decisions. Examples include medical diagnostics, criminal court sentencing, and pretrial release or detainment decisions (Novokmet et al. 2022).

An understudied research agenda pertains to how professionals interact with AI-based assessments when making decisions, particularly when their assessments diverge from those generated by AI technologies (Gerdon et al. 2022, Jussupow et al. 2022). Some investigations delve into aversion toward AI-based assessments, while others study overreliance on AI. Psychological research suggests that people's reluctance to rely on algorithmic advice stems from their perception of their own capabilities in comparison to the perceived accuracy of the algorithm. When individuals observe an algorithm making an error, they tend to become more confident in their ability to make the correct decision relative to their trust in the algorithm (Dietvorst et al. 2015, Kawakami et al. 2022).

Legal research examines the use of risk assessment tools in sentencing. Stevenson & Doleac (2022) conduct statistical analysis and uncover disparities between sentencing outcomes and recommendations provided by risk assessment tools. They argue that judges exercise discretion to mitigate adverse consequences of using risk assessment tools, which may come at the expense of reducing potential gains. Consequently, the adoption of risk assessment tools does not seem to yield discernible benefits in terms of public safety or reduced incarceration. Judges relied heavily on AI risk assessment tools during their initial adoption but gradually phased out their usage. Despite their valuable contribution, Stevenson and Doleac's methodology does not fully reveal how legal professionals use risk assessment tools.

Given that studies investigating this topic primarily rely on statistical analysis, simulation, or experiments, there is a need for qualitative research that employs observation and in-depth interviews. In this regard, two qualitative studies on how radiologists utilize diagnostic AI systems hold significant value in unraveling the complexities of the interactive decision-making process. Both studies find varying levels of engagement with AI systems among radiologists. Some radiologists try to make sense of AI assessments by relating their own knowledge to AI assessments and learn from this process, while others do not engage with AI systems. The latter group tends to find it difficult, time-consuming, and unproductive to engage with AI assessments due to their opacity, especially considering workload and time pressure (Jussupow et al. 2022, Lebovitz et al. 2022).

Jussupow et al. (2022) further explain the varying types of engagement. They found the use of AI systems by radiologists is linked to an individual radiologist's diagnostic self-efficacy, which is largely influenced by their experiences. Radiologists with moderate to low diagnostic self-efficacy tend to extensively utilize AI systems and make an effort to interpret AI assessments. Confirming AI assessments enhances their diagnostic confidence, while disconfirming assessments offers learning opportunities to avoid errors and gain insights into diagnostics or AI mistakes. Despite occasional incorrect assessments, they believe the benefits of AI outweigh the drawbacks. In comparison, radiologists with medium to high diagnostic self-efficacy perceive that AI systems do not influence their decisions. They attribute credibility to the AI system when their own assessments align with the AI assessments. They tend to dismiss conflicting advice without thorough deliberation, while sharing information about AI's errors with their colleagues. They perceive the system to be valuable only for less experienced colleagues. Radiologists with high diagnostic self-efficacy have minimal engagement with the AI system. They perceive the system as lacking value for themselves and their colleagues. Disconfirming advice from the AI system is seen as an unnecessary source of distraction. They emphasize the significance of personal skills, tacit knowledge, experience, and competence as crucial factors for accurate diagnostics. Consequently, radiologists with high diagnostic self-efficacy suggest removing the AI system from clinical practice.

Qualitative research investigating the experiences of social workers utilizing an AI-based decision tool to handle child welfare issues aligns with findings from studies involving radiologists. Social workers express frustration with the lack of intelligibility and transparency in the AI tool. Additionally, the assessments made by social workers often differ from AI-based assessments, as the latter do not consider crucial contextual knowledge and information. Consequently, most social workers hold negative evaluations of the system, affecting their engagement with it. The study also reveals that most social workers perceive AI-based assessments as playing a relatively minor, nondriving role in their overall decision-making processes (Kawakami et al. 2022). Overall, the evidence does not suggest that AI tools significantly augment the capabilities of social workers.

Despite the similarities between the studies on radiologists and social workers, there are critical differences concerning their autonomy and power within their organizations. In Kawakami et al.'s (2022) research, social workers exhibit a lower degree of autonomy in overriding AI-based assessments. Organizational pressure and incentives discourage social workers from disagreeing with these AI-based assessments, especially when they believe they already have an excessive number of overrides. As a result, they adopt AI-based assessments even when these assessments contradict their own best judgment. Moreover, the study highlights that social workers have limited power over the designers and administrators, who intentionally keep the model opaque to prevent social workers from gaming the system to produce desirable scores. Those in charge also largely overlook feedback from social workers regarding potential improvements. This lack of influence further exacerbates the challenges faced by social workers in effectively engaging with the AI tool.

EMBEDDING AUTOMATION AND AUGMENTATION INTO THE SOCIAL AND POLITICAL CONTEXTS

While the existing literature contributes to our understanding of the dangers, pitfalls, limitations, and benefits of robots and AI in automating tasks or augmenting human capabilities, there has been limited research focused on examining and explaining variations in these outcomes. We argue that future research should move in this direction and embed the study of automation and augmentation in concrete social and political contexts. Specifically, future research should examine how organizational and institutional factors, contexts, and mechanisms at different levels (e.g., firms, industries, subnations, countries, and beyond) influence diverse outcomes related to the development and utilization of robotics and AI. Sociologists are well-positioned to conduct such research, benefiting from the wealth of analytic tools available in various subfields of sociology.

Here, we propose two specific research agendas. First, as mentioned in the previous section, materiality, tacit knowledge, and expertise are recurring themes in existing studies concerning the potential and limitations of technology in automation and augmentation. Building on the work of Eyal & Pok (2011), we believe it will be fruitful to examine how the evolving configuration of expertise networks within and across organizations and fields influences the specific dangers, pitfalls, limitations, and benefits of technology developed to assist in task automation or labor augmentation. To do this, it is necessary to explore the interplay between professionals, semiprofessionals, technical experts, other type of workers, professional associations, organizations (e.g., government agencies, law firms, hospitals, tech firms, research institutes, labor unions, and nongovernmental organizations), the public, instruments (e.g., technological and legal instruments), forms of knowledge, competing claims of expertise and knowledge, professional identity, autonomy, and organizational and institutional environments. Such a research agenda will primarily focus on micro- and mesolevels of analysis and their interaction. Our approach aligns with that of AI and HCI communities in favor of adopting a multi-stakeholder perspective to study the design, development, and maintenance of AI (Delgado et al. 2021). However, it is situated in organizational and institutional analysis and dynamics and is more specific in terms of identifying relevant sociological concepts.

For example, one can extend the case study conducted by Kawakami et al. (2022) on social workers. The study reveals that social workers are granted limited autonomy to override AI assessments and have minimal influence on the development of the AI tool, despite voicing their concerns. The analytics team within the bureaucracy seems to prioritize maintaining a specific acceptable rate of overrides, thereby exerting power over social workers and their supervisors. Presumably, social workers in different organizations may have varying levels of autonomy and be embedded in distinct organizational structures and cultures. These organizational contexts can significantly influence the relationship and interaction between social workers, technical experts (e.g., AI model designers), and supervisors, leading to diverse modes of knowledge transmission, outcomes of competing claims of expertise, varied engagements with AI tools, and differing evaluations of the benefits of AI tools in enhancing social workers' capabilities.

As another illustration, one can compare the process and consequences of automation and augmentation in settings with varying power relations between workers and employers. Factories in both China and Germany have promoted the agenda of Industry 4.0. However, labor unions have a strong presence in Germany but not in China. Workers in Germany, in general, are likely to be able to protect themselves from the adverse consequences of automation and benefit from labor augmentation more than their counterparts in China (Haipeter 2020, Pfeiffer 2016).

Secondly, we also find it valuable to conduct more macrolevel analysis, focusing on the formation, change, and consequences of the larger institutional environment. Future research can examine different regimes that govern automation-related technology as well as the intended and unintended consequences of these regimes. Technology is not deterministic. As economist Jason Furman (2019) has pointed out, many high-income countries have experienced similar technological changes as the United States, yet the United States has had a greater increase in income inequality and higher overall levels of inequality than other high-income countries. Indeed, institutions and public policies can significantly influence whether and to what extent changes in technology shape socioeconomic outcomes and result in winners and losers.

To understand and examine the current moment, it is helpful to look back the history. In the 1960s, politicians, trade union leaders, scholars, and the general public in the United States grappled with the implications of new levels of automation made possible by computers and program-controlled machines. In 1964, activists, scholars, and technologists sent an open memorandum to President Lyndon B. Johnson, warning about the cybernation revolution and its potential for almost unlimited productivity with progressively less human labor. In response, President Johnson established the National Commission on Technology, Automation, and Economic Progress to address concerns about automation's impact on employment. The Commission aimed to harness technology's benefits for productivity and progress while safeguarding the interests of workers and families. National security was also a priority, given the Cold War context, as automation was considered vital to the country's security (Lei 2023).

Since the 1960s, however, conditions have changed significantly in the United States and globally. The New Deal order declined, replaced by the neoliberal order (Gerstle 2022). The Cold War ended, and neoliberal globalization emerged. Instead of addressing concerns in the 1960s, the United States witnessed the rise of a powerful and loosely regulated tech sector under a fragmented political and legal environment influenced by the Californian ideology. This ideology combines counterculture with a belief in the transformative power of new ITs, social liberalism, and economic liberalism (Turner 2006). Gradually, the United States saw the weakening of the welfare state and an increasingly poorly organized labor force. In this environment, technological changes did not result in broad-based prosperity but rather in increased inequality (Acemoglu & Johnson 2023).

Furthermore, the neoliberal globalization promoted by American politicians and multinational companies unexpectedly contributed to the rise of China as an economic and technological superpower, eventually leading to geopolitical tensions. In China, the government established a technological development regime characterized by a hyper-rational developmental state, an authoritarian political regime, and an asymmetrically symbiotic relationship between the government and tech companies. The dominant ideology blends high modernism, techno-nationalism, technological fetishism, and meritocracy. Similarly to the United States, China has a weak welfare state, a poorly organized labor force, and significant inequality. The government's efforts to facilitate high-tech-driven economic development have resulted in an instrument-centered rather than human-centered developmental regime, which pays little attention to marginalized groups impacted by technological and economic transformation (Lei 2023).

Today, the world witnesses the decline of neoliberal globalization and the coexistence of two superpowers—the United States and China—that compete for technological supremacy in the age of AI. Under these new conditions, the once hidden or disguised developmental state and industrial policy are back in the United States, as illustrated by the CHIPS and Science Act of 2022 in the United States (Lei 2023). The bill's spending is dedicated to chip manufacturing, research into AI, quantum computing, and robotics, among other areas.

In a review article, Wajcman (2017) questions which aspects are old and new in recent discussions about AI and robotics when we situate the recent wave of hype and fear historically. We consider the current moment a critical historical juncture. Compared with concerns in the past, issues related to automation and augmentation have grown more complex. In the past, the focus was solely on the threat to employment, rising inequality, and assisting workers in adapting through education and training. Now, new issues include bias, surveillance, threats to privacy, and threats to democracy. Additionally, the unfulfilled potential of technology, declining productivity growth, cybersecurity, data regulation, and competition policy are also areas of concern (Agrawal et al. 2019, Furman 2019). As geopolitical tensions rise, national security and technological self-reliance have become crucial, especially for superpowers like the United States and China (Lei 2023). Decisions and negotiations concerning these diverse issues, which could potentially conflict with each other, will significantly impact various aspects of life, including the potential for AI and robots to promote broad-based prosperity. Sociologists are well-positioned to study such political processes and their consequences.

CONCLUSION

The future of work agenda has remained relevant since the early 1970s, with the emergence of new technologies capable of task automation and labor augmentation. After reviewing the literature, we present the following observations and suggestions. Firstly, displacement effects from task automation persist. However, one should not assume an unequivocally increasing efficacy of technology in automation or augmentation, especially given the declining productivity growth in high-income countries and some large emerging economies in recent decades. Jobs with diverse tasks, requiring physical dexterity, flexibility, or tacit knowledge, and those protected by professional associations or trade unions, are less likely to be negatively impacted by automation-related technology. Despite the existence of countervailing effects, automation and augmentation could widen inequality between social groups, labor and capital, and firms without policy intervention, as better-off individuals and firms are more likely to mitigate the adverse consequences of automation and benefit from emerging technologies.

Secondly, despite the hype surrounding AI, the promise of AI in task automation and labor augmentation presents a mixed picture. AI tools can automate tasks in ways that cause harm, and even when no harms exist, users often express dissatisfaction or disengagement with AI-based assessments due to their opaqueness, errors, disregard for critical contexts, tacit knowledge, and domain expertise, as well as the demand for extra labor time and resources. These problems largely frustrate users who have to use AI at work, especially those with limited autonomy to override AI-based assessments from supervisors or technical experts.

Moving forward, we argue that sociological research should extend beyond identifying dangers, pitfalls, and limitations of robots and AI in automation and augmentation. Instead, future research should focus on specifying the conditions and mechanisms that ameliorate the adverse consequences of task automation and enhance labor augmentation. To achieve this, sociologists should systematically embed the study of automation and augmentation in concrete social and political contexts at multiple levels.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We presented this article at the 2023 Annual Meeting of the American Sociological Association. We extend our gratitude to the participants of our panel for their insightful questions and contributions.

LITERATURE CITED

Acemoglu D. 2021. *Harms of AI*. NBER Work. Pap. 9247

Acemoglu D, Autor D, Dorn D, Hanson GH, Price B. 2014. Return of the Solow paradox? IT, productivity, and employment in US manufacturing. *Am. Econ. Rev.* 104(5):394–99

Acemoglu D, Johnson S. 2023. *Power and Progress: Our Thousand-Year Struggle over Technology and Prosperity*. New York: PublicAffairs

Acemoglu D, Restrepo P. 2017. *Robots and jobs: evidence from US labor markets*. NBER Work. Pap. 23285

Acemoglu D, Restrepo P. 2018. The race between man and machine: implications of technology for growth, factor shares, and employment. *Am. Econ. Rev.* 108(6):1488–542

Acemoglu D, Restrepo P. 2019a. Artificial intelligence, automation, and work. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 197–236. Chicago: Univ. Chicago Press

Acemoglu D, Restrepo P. 2019b. Automation and new tasks: how technology changes labor demand. *J. Econ. Perspect.* 33(2):3–30

Acemoglu D, Restrepo P. 2020. The wrong kind of AI? Artificial intelligence and the future of labour demand. *Camb. J. Reg. Econ. Soc.* 13(1):25–35

Adler PS. 1992. Introduction. In *Technology and the Future of Work*, ed. PS Adler, pp. 3–14. Oxford, UK: Oxford Univ. Press

Agrawal A, Gans J, Goldfarb A. 2019. Introduction. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 1–20. Chicago: Univ. Chicago Press

Agrawal A, Gans J, Goldfarb A. 2023. Do we want less automation? *Science* 381(6654):155–58

Attard-Frost B, De los Ríos A, Walters DR. 2023. The ethics of AI business practices: a review of 47 AI ethics guidelines. *AI Ethics* 3(2):389–406

Attewell P. 1990. What is skill? *Work Occ.* 17(4):422–48

Attewell P. 1992. Skill and occupational changes in US manufacturing. In *Technology and the Future of Work*, ed. PS Adler, pp. 46–88. Oxford, UK: Oxford Univ. Press

Autor D. 2015. Why are there still so many jobs? The history and future of workplace automation. *J. Econ. Perspect.* 29(3):3–30

Autor D, Dorn D, Katz LF, Patterson C, Van Reenen J. 2020. The fall of the labor share and the rise of superstar firms. *Q. J. Econ.* 135(2):645–709

Autor D, Levy F, Murnane RJ. 2003. The skill content of recent technological change: an empirical exploration. *Q. J. Econ.* 118(4):1279–333

Autor D, Salomons A. 2018. *Is automation labor share-displacing? Productivity growth, employment, and the labor share*. Work. Pap., Brookings Inst., Washington, DC

Bailey S, Pierides D, Brisley A, Weisshaar C, Blakeman T. 2020. Dismembering organisation: the coordination of algorithmic work in healthcare. *Curr. Soc.* 68(4):546–71

Bainbridge L. 1983. Ironies of automation. *Automatica* 19(6):775–79

Balliester T, Elsheikhi A. 2018. *The future of work: a literature review*. Work. Pap. 29, ILO Res. Dep., Int. Labour Organ., Geneva

Bauer H, Brandl F, Lock C, Reinhart G. 2018. Integration of Industrie 4.0 in lean manufacturing learning factories. *Proc. Manuf.* 23:147–52

Benjamin R. 2019. Assessing risk, automating racism. *Science* 366(6464):421–22

Bijker WE. 2011. Tacit and explicit knowledge. *Technol. Cult.* 52(4):809–10

Brandt L, Litwack J, Mileva E, Wang L, Zhang W, Zhao L. 2020. *China's productivity slowdown and future growth potential*. Policy Res. Work. Pap. 9298, World Bank, Washington, DC

Brayne S, Christin A. 2020. Technologies of crime prediction: the reception of algorithms in policing and criminal courts. *Soc. Probl.* 68(3):608–24

Brooks C, Gherhes C, Vorley T. 2020. Artificial intelligence in the legal sector: pressures and challenges of transformation. *Camb. J. Reg. Econ. Soc.* 13(1):135–52

Broussard M. 2018. *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press

Broussard M. 2023. *More Than a Glitch: Confronting Race, Gender, and Ability Bias in Tech*. Cambridge, MA: MIT Press

Brynjolfsson E. 2022. The Turing trap: the promise and peril of human-like artificial intelligence. *Daedalus* 151(2):272–87

Brynjolfsson E, Rock D, Syverson C. 2019. Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 23–60. Chicago: Univ. Chicago Press

Buolamwini J, Gebru T. 2018. Gender shades: intersectional accuracy disparities in commercial gender classification. *Proc. Mach. Learn. Res.* 81:77–91

Burrell J, Fourcade M. 2021. The society of algorithms. *Annu. Rev. Sociol.* 47:213–37

Chen Y, Stavropoulou C, Narasinkan R, Baker A, Scarbrough H. 2021. Professionals' responses to the introduction of AI innovations in radiology and their implications for future adoption: a qualitative study. *BMC Health Serv. Res.* 21:813

Cheng H, Jia R, Li D, Li H. 2019. The rise of robots in China. *J. Econ. Perspect.* 33(2):71–88

Chiacchio F, Petropoulos G, Pichler D. 2018. *The impact of industrial robots on EU employment and wages: a local labour market approach*. Work. Pap. 2018/02, Breugel, Brussels

Chin J, Lin L. 2022. *Surveillance State: Inside China's Quest to Launch a New Era of Social Control*. New York: St. Martin's

Ciuriak D. 2018. *The economics of data: implications for the data-driven economy*. Work. Pap., Cent. Int. Gov. Innov., Waterloo, ON, Can.

Clifton J, Glasmeier A, Gray M. 2020. When machines think for us: the consequences for work and place. *Camb. J. Reg. Econ. Soc.* 13(1):3–23

Collins HM. 2010. *Tacit and Explicit Knowledge*. Chicago: Univ. Chicago Press

Dahlin E. 2019. Are robots stealing our jobs? *Socius* 5:1–14

Dao MC, Das M, Koczan Z, Lian W. 2017a. Understanding the downward trend in labor income shares. In *World Economic Outlook, April 2017: Gaining Momentum?*, pp. 121–72. Washington, DC: Int. Monetary Fund

Dao MC, Das M, Koczan Z, Lian W. 2017b. *Why Is Labor Receiving a Smaller Share of Global Income? Theory and Empirical Evidence*. Washington, DC: Int. Monetary Fund

Dargan S, Kumar M, Ayyagari MR, Kumar G. 2020. A survey of deep learning and its applications: a new paradigm to machine learning. *Arch. Comput. Methods Eng.* 27:1071–92

De Bruyn A, Viswanathan V, Beh YS, Brock JK, Von Wangenheim F. 2020. Artificial intelligence and marketing: pitfalls and opportunities. *J. Interact. Mark.* 51(1):91–105

Delgado F, Yang S, Madaio M, Yang Q. 2021. Stakeholder participation in AI: beyond 'add diverse stakeholders and stir.' arXiv:2111.01122 [cs.AI]

Dietvorst BJ, Simmons JP, Massey C. 2015. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *J. Exp. Psychol. Gen.* 144(1):114–26

Dorschel R. 2021. Discovering needs for digital capitalism: the hybrid profession of data science. *Big Data Soc.* 8(2). **https://doi.org/10.1177/20539517211040760**

Dressel J, Farid H. 2018. The accuracy, fairness, and limits of predicting recidivism. *Sci. Adv.* 4: eaao5580

Ensmenger N. 2010. *The Computer Boys Take Over: Computers, Programmers, and the Politics of Technical Expertise*. Cambridge, MA: MIT Press

Epstein S. 1995. The construction of lay expertise: AIDS activism and the forging of credibility in the reform of clinical trials. *Sci. Technol. Hum. Values* 20(4):408–37

Eubanks V. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's

Eyal G, Pok G. 2011. *From a sociology of professions to a sociology of expertise*. Paper presented at CAST Workshop on Security Expertise, Univ. Copenhagen, Copenhagen, Den., June 15–17

Ferrer X, van Nuenen T, Such JM, Coté M, Criado N. 2021. Bias and discrimination in AI: a cross-disciplinary perspective. *IEEE Technol. Soc. Mag.* 40(2):72–80

Furman J. 2019. Should we be reassured if automation in the future looks like automation in the past? In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 317–28. Chicago: Univ. Chicago Press

Gerdon F, Bach RL, Kern C, Kreuter F. 2022. Social impacts of algorithmic decision-making: a research agenda for the social sciences. *Big Dat. Soc.* 9(1). **https://doi.org/10.1177/20539517221089305**

Gerstle G. 2022. *The Rise and Fall of the Neoliberal Order: America and the World in the Free Market Era*. Oxford, UK: Oxford Univ. Press

Goos M, Manning A. 2007. Lousy and lovely jobs: the rising polarization of work in Britain. *Rev. Econ. Stat.* 89(1):118–33

Gordon RJ. 2016. *The Rise and Fall of American Growth: The US Standard of Living Since the Civil War*. Princeton, NJ: Princeton Univ. Press

Gray ML, Suri S. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston: Houghton Mifflin Harcourt

Green B. 2021. The contestation of tech ethics: a sociotechnical approach to technology ethics in practice. *J. Soc. Comput.* 2(3):209–25

Green B, Chen Y. 2019. Disparate interactions: an algorithm-in-the-loop analysis of fairness in risk assessments. In *FAT* '19: Proceedings of the Conference on Fairness, Accountability, and Transparency*, pp. 90–99. New York: ACM

Greene D, Hoffmann AL, Stark L. 2019. Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, ed. TX Bui, pp. 2122–31. Honolulu, HI: Univ. Hawai'i at Mānoa

Griesbach K, Reich A, Elliott-Negri L, Milkman R. 2019. Algorithmic control in platform food delivery work. *Socius* 5. **https://doi.org/10.1177/2378023119870041**

Haipeter T. 2020. Digitalisation, unions and participation: the German case of 'Industry 4.0.' *Ind. Relat. J.* 51(3):242–60

Joyce K, Smith-Doerr L, Alegria S, Bell S, Cruz T, Hoffman SG, et al. 2021. Toward a sociology of artificial intelligence: a call for research on inequalities and structural change. *Socius* 7. **https://doi.org/10.1177/2378023121999581**

Jussupow E, Spohrer K, Heinzl A. 2022. Radiologists' usage of diagnostic AI systems. *Bus. Inf. Syst. Eng.* 64(3):293–309

Jussupow E, Spohrer K, Heinzl A, Gawlitza J. 2021. Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence. *Inf. Syst. Res.* 32(3):713–35

Katz LF, Murphy KM. 1992. Changes in relative wages, 1963–1987: supply and demand factors. *Q. J. Econ.* 107(1):35–78

Kawakami A, Sivaraman V, Cheng HF, Stapleton L, Cheng Y, et al. 2022. Improving human-AI partnerships in child welfare: understanding worker practices, challenges, and desires for algorithmic decision support. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. New York: ACM

Kellogg K, Valentine MA, Christin A. 2020. Algorithms at work: the new contested terrain of control. *Acad. Manag. Ann.* 14(1):366–410

Kluttz DN, Mulligan DK. 2019. Automated decision support technologies and the legal profession. *Berkeley Technol. Law J.* 34(3):853–90

Kshetri N. 2021. Data labeling for the artificial intelligence industry: economic impacts in developing countries. *IT Prof.* 23(2):96–99

Lam A. 2000. Tacit knowledge, organizational learning and societal institutions: an integrated framework. *Organ. Stud.* 21(3):487–513

Lebovitz S, Lifshitz-Assaf H, Levina N. 2022. To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis. *Organ. Sci.* 33(1):126–48

Lei YW. 2021. Delivering solidarity: platform architecture and collective contention in China's platform economy. *Am. Sociol. Rev.* 86(2):279–309

Lei YW. 2022. Upgrading China through automation: manufacturers, workers and the techno-developmental state. *Work Employ. Soc.* 36(6):1078–96

Lei YW. 2023. *The Gilded Cage: Technology, Development, and State Capitalism in China*. Princeton, NJ: Princeton Univ. Press

Madaio M, Egede L, Subramonyam H, Vaughan JW, Wallach H. 2022. Assessing the fairness of AI systems: AI practitioners' processes, challenges, and needs for support. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 6. New York: ACM

Mateescu A, Elish M. 2019. *AI in context: the labor of integrating new technologies*. Rep., Data Soc. Res. Inst., New York. **https://apo.org.au/node/217456**

McCarthy J. 2007. *What is artificial intelligence?* Work. Pap., Comput. Sci. Dep., Stanford Univ., Stanford, CA. **https://www.diochnos.com/about/McCarthyWhatisAI.pdf**

Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. 2021. A survey on bias and fairness in machine learning. *ACM Comput. Surv.* 54(6):115

Metcalf J, Moss E. 2019. Owning ethics: corporate logics, Silicon Valley, and the institutionalization of ethics. *Soc. Res. Int. Q.* 86(2):449–76

Mokyr J. 2014. Secular stagnation? Not in your life. In *Secular Stagnation: Facts, Causes and Cures*, ed. R Baldwin, C Teulings, pp. 83–90. London: Cent. Econ. Policy Res.

Munn L. 2022a. *Automation Is a Myth*. Stanford, CA: Stanford Univ. Press

Munn L. 2022b. The uselessness of AI ethics. *AI Ethics* 3:869–77

Noble DF. 1984. *Forces of Production: A Social History of Industrial Automation*. New York: Random House

Noble SU. 2018. *Algorithms of Oppression*. New York: NYU Press

Novokmet A, Tomičić Z, Vinković Z. 2022. Pretrial risk assessment instruments in the US criminal justice system—what lessons can be learned for the European Union. *Int. J. Law Inform. Tech.* 30(1):1–22

Obermeyer Z, Powers B, Vogeli C, Mullainathan S. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366(6464):447–53

Pardo-Guerra JP. 2019. *Automating Finance: Infrastructures, Engineers, and the Making of Electronic Markets*. Cambridge, UK: Cambridge Univ. Press

Pfeiffer S. 2016. Robots, Industry 4.0 and humans, or why assembly work is more than routine work. *Societies* 6(2):1–26

Polanyi M. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: Univ. Chicago Press

Polanyi M. 1967. *The Tacit Dimension*. Garden City, NY: Anchor

Raj M, Seamans R. 2019. Artificial intelligence, labor, productivity, and the need for firm-level data. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 553–66. Chicago: Univ. Chicago Press

Rodgers I, Armour J, Sako M. 2023. How technology is (or is not) transforming law firms. *Annu. Rev. Law Soc. Sci.* 19:299–317

Rosenblat A. 2018. *Uberland: How Algorithms Are Rewriting the Rules of Work*. Berkeley: Univ. Calif. Press

RSNA (Radiol. Soc. N. Am.). 2022. Radiology facing a global shortage. *RSNA News Blog*, May 10. **https://www.rsna.org/news/2022/may/global-radiologist-shortage**

Sako M, Qian M, Attolini J. 2022. Future of professional work: evidence from legal jobs in Britain and the United States. *J. Prof. Organ.* 9(2):143–69

Schuetz PNK. 2021. Fly in the face of bias: algorithmic bias in law enforcement's facial recognition technology and the need for an adaptive legal framework. *Law Inequal*. 39:221–54

Shen J, Zhang CJP, Jiang B, Chen J, Song J, et al. 2019. Artificial intelligence versus clinicians in disease diagnosis: systematic review. *JMIR Med. Inform.* 7(3):e10010

Shestakofsky B. 2017. Working algorithms: software automation and the future of work. *Work Occup*. 44(4):376–423

SIPAC (Suzhou Ind. Park Adm. Comm.). 2021. *From 2019 to 2021, the Ministry of Human Resources and Social Security released 4 batches of 56 new occupations for skilled personnel (continuously updated)*. Fact Sheet, Suzhou Ind. Park Adm. Comm., Suzhou, China. **http://www.sipac.gov.cn/ldhshbzj/tzgg/202105/9a5f774b821e4289a66aa9e992c3f0cb.shtml**

Spencer D, Slater G. 2020. No automation please, we're British: technology and the prospects for work. *Camb. J. Reg. Econ. Soc.* 13(1):117–34

Spenner KI. 1990. Skill: meanings, methods, and measures. *Work Occup*. 17(4):399–421

Stevenson MT, Doleac JL. 2022. *Algorithmic risk assessment in the hands of humans*. IZA Discuss. Pap. 12853, Inst. Labor Econ., Bonn, Ger.

Susskind RE, Susskind D. 2015. *The Future of the Professions: How Technology Will Transform the Work of Human Experts*. Oxford, UK: Oxford Univ. Press

Syverson C. 2017. Challenges to mismeasurement explanations for the US productivity slowdown. *J. Econ. Perspect.* 31(2):16–86

Tassinari A, Maccarrone V. 2020. Riders on the storm: workplace solidarity among gig economy couriers in Italy and the UK. *Work Employ. Soc.* 34(1):35–54

Trajtenberg M. 2019. Artificial intelligence as the next GPT: a political-economy perspective. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 175–86. Chicago: Univ. Chicago Press

Turner R. 2006. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago: Univ. Chicago Press

Valentine M, Hinds R. 2022. *How algorithms change occupational expertise by prompting explicit articulation and testing of experts' theories*. Work. Pap. 2022, Stanford Manag. Sci. Eng., Stanford, CA

Vallas SP. 2017. Platform capitalism: What is at stake for workers? *New Labor Forum* 28(1):48–59

Van Reenen J. 2011. Wage inequality, technology and trade: 21st-century evidence. *Labour Econ*. 18(6):730–41

Vogt G, König ASL. 2021. Robotic devices and ICT in long-term care in Japan: their potential and limitations from a workplace perspective. *Contemp. Jpn.* 35:270–90

Wajcman J. 2017. Automation: Is it really different this time? *Br. J. Sociol.* 68(1):119–27

Zuboff S. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs