Toward understanding the impact of artificial intelligence on labor

Morgan R. Frank*, David Autor†, James E. Bessen*, Erik Brynjolfsson†, Manuel Cebrian*, David J. Deming†,§, Maryann Feldman†, Matthew Groh†, José Lobo, Esteban Moro†, Dashun Wang†, Hyejin Youn†, and Iyad Rahwan*§‡,§,k

Edited by Jose A. Scheinkman, Columbia University, New York, NY, and approved February 28, 2019 (received for review January 18, 2019)

Rapid advances in artificial intelligence (AI) and automation technologies have the potential to significantly disrupt labor markets. While AI and automation can augment the productivity of some workers, they can replace the work done by others and will likely transform almost all occupations at least to some degree. Rising automation is happening in a period of growing economic inequality, raising fears of mass technological unemployment and a renewed call for policy efforts to address the consequences of technological change. In this paper we discuss the barriers that inhibit scientists from measuring the effects of AI and automation on the future of work. These barriers include the lack of high-quality data about the nature of work (e.g., the dynamic requirements of occupations), lack of empirically informed models of key microlevel processes (e.g., skill substitution and human–machine complementarity), and insufficient understanding of how cognitive technologies interact with broader economic dynamics and institutional mechanisms (e.g., urban migration and international trade policy). Overcoming these barriers requires improvements in the longitudinal and spatial resolution of data, as well as refinements to data on workplace skills. These improvements will enable multidisciplinary research to quantitatively monitor and predict the complex evolution of work in tandem with technological progress. Finally, given the fundamental uncertainty in predicting technological change, we recommend developing a decision framework that focuses on resilience to unexpected scenarios in addition to general equilibrium behavior.

Artificial Intelligence (AI) is a rapidly advancing form of technology with the potential to drastically reshape US employment (1, 2). Unlike previous technologies, examples of AI have applications in a variety of highly educated, well-paid, and predominantly urban industries (3), including medicine (4, 5), finance (6), and information technology (7). With AI’s potential to change the nature of work, how can policy makers facilitate the next generation of employment opportunities? Studying this question is made difficult by the complexity of economic systems and AI’s differential impact on different types of labor.

While technology generally increases productivity, AI may diminish some of today’s valuable employment opportunities. Consequently, researchers and policy makers worry about the future of work in both advanced and developing economies worldwide. As an example, China is making AI-driven technology the centerpiece of its economic development plan (8). Automation concerns are not new to AI, and examples date back even to the advent of written language. In ancient Greece (ca. 370 BC), Plato’s Phaedrus (9) described how writing would displace human memory...
and reading would substitute true knowledge with mere data. More commonly, historians point to the Industrial Revolution and the riots of 19th-century Luddites (10) as examples where technological advancement led to social unrest. Two examples from the recent past echo these concerns.

Vasily Leontief, winner of the 1973 Nobel Prize in Economics, noted in 1952, “Labor will become less and less important... More workers will be replaced by machines. I do not see that new industries can employ everybody who wants a job” (11).

Similarly, US Attorney General Robert F. Kennedy commented in 1964, “Automation provides us with wondrous increases of production and information, but does it tell us what to do with the men the machines displace? Modern industry gives us the capacity for unparalleled wealth—but where is our capacity to make that wealth meaningful to the poor of every nation?” (12).

However, despite these long-lasting and oft-recurring concerns, society underwent profound transformations, the economy continued to grow, technology continues to advance, and workers continue to have jobs. Given this history of concern, what makes human labor resilient to automation? Is AI a fundamentally new concern from technologies of the past?

Answering these questions requires a detailed knowledge that connects AI to today’s workplace skills. Each specific technology alters the demand for specific types of labor, and thus the varying skill requirements of different job titles can obfuscate technology’s impact. In general, depending on the nature of the job, a worker may be augmented by technology or in competition with it (13–15). For example, technological advancements in robotics can diminish wages and employment opportunities for manufacturing workers (16, 17). However, technological change does not necessarily produce unemployment, and, in the case of AI, cognitive technology may actually augment workers (18, 19). For instance, machine learning appears to bolster the productivity of software developers while also creating new investment and manufacturing opportunities (e.g., autonomous vehicles). Complicating matters further, the skill requirements of occupations do not remain static, but instead change with changing technology (19, 20).

In the remainder of this article, we describe how these competing dynamics combined with insufficient data might allow contrasting perspectives to coexist. In particular, we argue that the limitations into data about workplace tasks and skills restricts the viable approaches to the problem of technological change and the future of work. We offer suggestions to improve data collection with the goal of enriching models for workplace skills, employment, and the impact of AI. Finally, we suggest insights that improved data could provide in combination with a methodological focus on resilience and forecasting.

Contrasting Perspectives

Doomsayer’s Perspective. Technology improves to make human labor more efficient, but large improvements may yield deleterious effects for employment. This obsolescence through labor substitution leads many to worry about “technological unemployment” and motivates efforts to forecast AI’s impact of jobs. One study assessed recent developments in AI to conclude that 47% of current US employment is at high risk of computerization (23), while a contrasting study, using a different methodology, concluded that a less alarming 9% of employment is at risk (24). Similar studies have looked at the impact of automation on employment in other countries and reached sobering conclusions: Automation will affect 35% of employment in Finland (25), 59% of employment in Germany (26), and 45 to 60% of employment across Europe (27).

Critics have complained that prospective studies lack validation, but retrospective studies also find that robotics are diminishing employment opportunities in US manufacturing (17, 28)[although not in Germany (29)].

Optimist’s Perspective. Optimists suggest that technology may substitute for some types of labor but that efficiency gains from technological augmentation outweigh transition costs (30–34) and, in many cases, technology increases employment for workers who are in in direct competition with it (19, 35) [although recent follow-up work suggests these are temporary gains (28)]. Furthermore, the skill requirements of each job title are not static and actually evolve over time to reflect evolving labor needs. For example, workers may require more social skills because those skills remain difficult to automate (20). Even if technology depresses employment for some types of labor, it can create new needs and new opportunities through “creative destruction” (36–38). For instance, the replacement of equestrian travel with automobiles spurred demand for new roadside amenities, such as motels, gas stations, and fast food (39).

Unifying Perspectives. On one hand, multiple dynamics accompany technological change and create uncertainty about the future of work. On the other hand, experts agree that occupations are best understood as abstract bundles of skills (18, 40) and that technology directly impacts demand for specific skills instead of acting on whole occupations all at once (16, 19, 35, 41). Therefore, a detailed framework that connects specific skill types to career mobility (18, 42) and to whole urban workforces (40) may help to unify competing perspectives (Fig. 1C). Existing studies have argued theoretically that different skill types underpin aggregate labor trends, such as job polarization (16) and urban migration (43, 44), but robust empirical validation is made difficult by the sparsity of modern skills data and their temporal sparsity.

Overcoming Barriers to Forecasting the Future of Work

In this section we identify barriers to our scientific modeling of technological change and the future of work. Along with each barrier, we propose a potential solution that could enable improvement in forecasting labor trends. We provide a summary of these barriers and solutions in Table 1.

Barrier: Sparse Skills Data. Forecasting automation from AI requires skills data that keep pace with rapidly advancing technology [e.g., Moore’s Law (45), robots in manufacturing (17), and patent production (46–48)]. While skill types inform the theory of labor and technological change (1, 18, 21, 49), standard labor data focus on aggregate statistics, such as wage and employment numbers, and can lack resolution into the specifics that distinguish different job titles and different types of work. For example, previous studies have empirically observed a “hollowing” of the middle-skill jobs described by increasing employment share for low-skill and high-skill occupations at the expense of middle-skill occupations (16, 35) (reproduced in Fig. 1A). These studies use skills to explain labor trends but are limited empirically to measuring annual wages instead of skill content directly. While wages may correlate with specific skills, wage alone fails to capture the defining features of an occupation, and models focused on only cognitive and physical labor fail to explain responses to technological change (21).

As another approach, data on educational requirements can add resolution to employment trends (50–52). For instance, jobs that require a bachelor’s degree may identify cognitive workers...
who are less susceptible to automation. Ideally, educational institutions train workers to possess valuable skills that lead to higher wages (53). However, looking at education and wages alone has proven insufficient to explain stagnating returns on education (16, 54, 55) and slow wage growth despite increases in national productivity (14, 15, 41) (Fig. 18).

Improving data on the skills required to perform specific job tasks may provide better insights than wages and education alone. For example, previous studies have considered occupations as routine or nonroutine and cognitive or physical (21, 56–63) or looked at specific types of skills in relation to augmentation and substitution from technology (18, 41). Increasing a labor model’s specificity into workplace tasks and skills might further resolve labor trends and improve predictions of automation from AI. As an example, consider that civil engineers and medical doctors are both high-wage, cognitive, nonroutine occupations affected by technology (3, 65). However, employment trends and changing demand for specific workplace skills, thus altering the connections between skill pairs. As an example, machine vision software may impact the demand for human labor for some visual task. These alterations can accumulate and diffuse throughout the entire system as aggregate labor trends described in A and B.

Frank et al.
they are installing things on the job (see Fig. 2A and SI Appendix, section 1 for calculation). More generally, any static taxonomy for workplace skills is not ideal for a changing economy: Should mathematics and programming be two separate workplace skills given that they are both computational? Conversely, is “programming” too broad given the variety of existing software and programming languages? Perhaps it is more appropriate to specify programming tasks or specific programming languages (see Fig. 2B for an example), especially given the rapid development of AI and machine learning. Likely, the correct abstraction is situation-dependent, but O*NET data offer limited flexibility.

Table 1. Tabulating the current barriers to forecasting the future of work along with proposed solutions

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Potential solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sparse skills data</td>
<td>- Adaptive skill taxonomies</td>
</tr>
<tr>
<td></td>
<td>- Connect susceptible skills to new technology</td>
</tr>
<tr>
<td></td>
<td>- Improve temporal resolution of data collection</td>
</tr>
<tr>
<td></td>
<td>- Use data from career web platforms</td>
</tr>
<tr>
<td>Limited modeling of resilience</td>
<td>- Explore out-of-equilibrium dynamics</td>
</tr>
<tr>
<td></td>
<td>- Identify workplace skill interdependencies</td>
</tr>
<tr>
<td></td>
<td>- Connect skill relationships to worker mobility</td>
</tr>
<tr>
<td></td>
<td>- Relate worker mobility to economic resilience in cities</td>
</tr>
<tr>
<td></td>
<td>- Explore models of resilience from other academic domains</td>
</tr>
<tr>
<td>Places in isolation</td>
<td>- Labor dependencies between places (e.g., cities)</td>
</tr>
<tr>
<td></td>
<td>- Identify skill sets of local economies</td>
</tr>
<tr>
<td></td>
<td>- Identify heterogeneous impact of technology across places</td>
</tr>
<tr>
<td></td>
<td>- Use intercity connections to study national economic resilience</td>
</tr>
</tbody>
</table>

Granular skills data will help elucidate the micro-scale impact of AI and other technologies in labor systems. For instance, the specifications of recent patents might suggest automatable types of labor in the near future (46–48), thus elucidating the impact of technological change at the granularity of workplace-specific tasks and skills. The distribution of skill categories within occupations and over individuals’ careers can reveal how occupational skill requirements evolve. As an example, consider that occupations such as software developer dynamically change the skill requirements in job listings (e.g., “programming” in the 1990s vs. “Python,” “Java,” “Kubernetes,” etc. today) to reflect the tools

Fig. 2. Since the skill requirements of occupations may inform opportunities for career mobility, abstract skill data may obfuscate important labor trends. (A) We use O*NET data to identify the characteristic skill requirements for truck drivers, plumbers, and software developers (see SI Appendix, section 1 for calculation). Individual skills may be unique to an occupation (e.g., operating vehicles) or shared between occupations (e.g., low-light vision). The skill of installation is required by both plumbers and software developers, but this skill may not mean the same thing to workers in these two occupations. Programming is a skill required by software developers, but the coarseness of this skill definition may hide important dynamics brought on by new technology, including AI. (B) For example, we provide the percentage of Google searches for coding tutorials by programming language. Trends are smoothed using locally weighted scatter plot smoothing (see SI Appendix, section 2 for calculation). The Python programming language is widespread in the field of machine learning. Therefore, the increased ubiquity of AI and, in particular, machine learning may contribute to Python’s steady growth in popularity.
and required specialization of the time. Understanding the dynamics of specific skills combined with the incomes within occupations can capture the marginal value of different skills despite the dynamic nature of work.

Online career platforms offer an example of the empirical possibilities facilitated by nontraditional and new data sources. These websites collect real-time data that reflect labor dynamics in certain industries. Data from workers’ resumes can improve our understanding of education and careers, as well as identifying workers’ transitions between occupations and skill sets. Additionally, job postings capture fluctuations in labor demands and demonstrate changes in demand for specific skills. Combined, these two sources of skills data offer an adaptive granular view into the changing nature of work that may detail where labor disconnects exist. Access to these private data sources is currently restricted and typically requires a data-sharing agreement that protects personally identifiable information and other proprietary information. Of course, personal privacy and issues of representative sampling are inherent to these data, but increased access could meaningfully augment currently available open data on employment and workplace skills. One potential solution is to construct a secure environment for the sharing of detailed skills and career data that is similar to the recent Social Science One partnership (69) (see https://socialscience.one).

Barrier: Limited Modeling of Resilience. Recent studies show that historical technology-driven trends may not capture the AI-driven trends we face today. Consequently, some have concluded that AI is a fundamentally new technology (3, 65). If the trends of the past are not predictive of the employment trends from current or future technologies, then how can policy makers maintain and create new employment opportunities in the face of AI? What features of a labor market lead to generalized labor resilience to technological change?

It is difficult to construct resilient labor markets because of the uncertainty around technology’s impact on labor. For instance, designing viable worker retraining programs requires detailed knowledge of the local workforce, fluency with current technology, and an understanding of the complex dependencies between regional labor markets around the world (70, 71). Technology typically performs specific tasks and may alter demand for specific workplace skills as a result. These micro-scale changes to skill demand can accumulate into systemic labor trends including occupational skill redefinition, employment redistribution (e.g., job creation and technological unemployment), and geographic redistribution (e.g., worker migration). Forecasting these complex effects requires a detailed understanding of the pathways along which these dynamics occur.

As an emblematic example of these complex dynamics, consider the competition between human bank tellers and automated teller machines (ATMs) (described in ref. 72). Unexpectedly, national employment for bank tellers rose with the adoption of ATMs. One explanation is demand elasticity: As ATMs decreased the operating cost of bank branches, more bank branches opened nationwide to meet rising consumer demand. Another more complicated reason is the accompanying shift in fundamental skill requirements from clerical ability to social and persuasive skills used by salespeople and customer service representatives. The story of bank tellers and ATMs is only fully captured by connecting the job-level changes in occupational skill composition with the system-level dynamics of demand brought on by increased efficiency. Accordingly, an updated framework for labor and AI must capture the interactions of microscopic workplace skills in combination to produce macroscopic labor trends, such as employment shifts, job polarization, and workers’ spatial mobility (for example, see Fig. 3B).

Existing theory of the matching process between job seekers and job vacancies (22) provides a stylized description of the matching process that lacks resolution into skill demand. Mapping
the space of skill interdependencies (e.g., Fig. 1C) could inform training and job assistance programs by identifying which types of work—and which locations—may experience augmentation and/or substitution with new technology. The detailed skill requirements of occupations determine the career mobility of individual workers, and thus changes to the demand for certain skills have the potential to redefine viable career trajectories and worker flow between occupations (e.g., middle layer of Fig. 1D). Therefore, mapping the relationships between jobs and skills that produce employment opportunities is a vital step for policy makers in the face of technological change.

In related domains, tools from network science have already provided new insights into modeling (and minimizing) systemic risk (73) in global credit (74) and financial industries (75), forecasting the future exports of national economies (76–78), mapping worker flows between industries (79) and firms (80), and charting the changing industrial composition of cities (81–83) and municipalities (84). Therefore, identifying the pathways along which labor dynamics (e.g., how skills determine workers’ career mobility) occur may provide similarly useful insights into the impact of AI on labor. Similar methods have been used to measure ecological resilience based on the structure of mutualistic interspecies interactions (66, 85). These methods often rely on the size and density of interconnected entities to estimate systemic resilience to species removal—perhaps analogous to diminishing demand for a skill with new technology (e.g., Fig. 3A).

Mapping skill dependencies will require appropriate data-handling methods. The ideal skills data should reflect the dynamic nature of skill representation, and so the methods we use to detect, categorize, and measure the demand for skills must be adaptable as well. Perhaps ironically, advanced AI techniques may help. Tools from machine learning (ML) and natural language processing (NLP) may capture the latent structure in complex high-dimensional data, thus making them ideal tools for the proposed application (and other applications in econometrics (86)). For example, NLP may be used to process historical skills data from the Dictionary of Occupational Titles into a format akin to the modern O*NET data. ML can be used on longitudinal job postings data to identify trends in skill demands that may reflect changes in technological ability. Combining these modern computational methods with relevant sources of data may foster new insights into labor dynamics at a high temporal resolution. In turn, these methodological improvements can bolster labor forecasts and policy makers’ ability to respond to real-time labor trends.

Barrier: Places in Isolation. The impact of AI and automation will vary greatly across geography, which has implications for the labor force, urban–rural discrepancies, and changes in the income distribution (87). The study of AI and automation are largely focused on national employment trends and national wealth disparity. However, recent work demonstrates that some places (e.g., cities) are more susceptible to technological change than others (17, 64). Occupations form a network of dependencies which constrain how easily jobs can be replaced by technology (82, 88). Therefore, the health of the aggregate labor market may depend on the impact of technology on specific urban and rural labor markets (73, 84).

Although technological change alters demand for specific workplace tasks and skills, current skills data mask the specific skill sets that comprise and differentiate the workforces of different

Fig. 4. A data pipeline that overcomes barriers to studying the future of work. (A) Inputs into the data pipeline include structured and unstructured data that detail regional variations in labor and granular skills data in relation to technological change. (B) Data from a variety of sources will need to be centralized and processed into a form that economists and data scientists can easily use (e.g., NLP to identify skill from resume and job postings). (C) Cleaned data feed a model for both the intercity (e.g., worker migration) and intracity (e.g., changes to local career mobility) labor trends brought on by technological change. (D) Outputs from this model will forecast the labor impact of technological change. These forecasts will inform policy makers seeking to implement prudent policy and individual workers attempting to navigate their careers.
geographies. In part, this is because skills data from nationwide surveys, such as the O*NET database, average over the regional variability in the required skills of workers with shared job titles. For example, software developers seeking employment in Silicon Valley may need to advertise more specific skill sets than similar employees in a shallower labor market (following the division of labor theory). Exacerbating this trend, the same AI technologies that augment high-wage cognitive employment are more abundant in large cities, while the physical low-wage tasks that are most readily replaced by robotics are more abundant in small cities and rural communities. This observation suggests that national wealth disparity is reflected in the wealth disparity between large and small cities akin to wage inequality across individuals.

Improved models for spatial interdependencies require more granular skills data (discussed above) and new insights into the mechanisms that create today’s cross-sectional geographic trends. For instance, how do university towns, where people gain valuable cognitive skills, contribute to the productivity of large cities? Do these economic connections help explain why university towns perform surprisingly well compared with similarly sized cities according to socioeconomic indicators [including exposure to automation (64)]?

Furthermore, just as internal connectivity determines urban economic resilience (83), so too can the connections between US cities underpin the economic health of the national economy (48). For instance, an interruption in the supply chain of well-educated cognitive workers may stifle an urban economy that normally attracts skilled workers. Therefore, it behooves policy makers to understand the connections between their local labor market and other urban labor markets to assess the resilience of their local economy. Since employment opportunities are central in people’s decision to relocate (43) and skill matching is essential to the job matching process (22), understanding the constituent skill sets in cities can inform models for the spatial mobility of workers and improve our understanding of career mobility and career incentives.

Conclusion

AI has the potential to reshape skill demands, career opportunities, and the distribution of workers among industries and occupations in the United States and in other developed and developing countries. However, researchers and policy makers are underequipped to forecast the labor trends resulting from specific cognitive technologies, such as AI. Typically, technology is designed to perform a specific task which alters demand for specific workplace skills. The resulting alterations to skill demands diffuse throughout the economy, influencing occupational skill requirements, career mobility, and societal well-being (e.g., impacts to workers’ social identity). Identifying the specific pathways of these dynamics has been constrained by coarse historical data and limited tools for modeling resilience. We can overcome these obstacles, however, by prioritizing data collection that is detailed, responsive to real-time changes in the labor market, and respects regional variability (see Fig. 4 for a data-pipeline schematic). Specifically, better access to unstructured data from resumes and job postings along with new indicators for recent technological change (e.g., patent data) and models for both intercity and intracity labor dependencies will enable new and promising techniques for understanding and forecasting the future of work. This improved data collection will enable the use of new data-driven tools, including machine learning applications and systemic modeling that more accurately reflects the complexity of labor systems. New data will lead to new research that enriches our understanding of the impact of technology on modern labor markets.

Acknowledgments

This work summarizes insights from the workshop on Innovation, Cities, and the Future of Work, which was funded by NSF Grant 1733545. This work was supported by the Massachusetts Institute of Technology (MIT) and the MIT Initiative on the Digital Economy.

4 Barlow M (2016) AI and Medicine (O’Reilly Media, Sebastopol, CA).
10 Jones SE (2013) Against Technology: From the Luddites to Neo-Luddism (Routledge, Abingdon, UK).


