The Wrong Kind of AI? Artificial Intelligence and the Future of Labor Demand*

Daron Acemoglu       Pascual Restrepo
MIT                  Boston University

March 5, 2019

Abstract

Artificial Intelligence is set to influence every aspect of our lives, not least the way production is organized. AI, as a technology platform, can automate tasks previously performed by labor or create new tasks and activities in which humans can be productively employed. Recent technological change has been biased towards automation, with insufficient focus on creating new tasks where labor can be productively employed. The consequences of this choice have been stagnating labor demand, declining labor share in national income, rising inequality and lower productivity growth. The current tendency is to develop AI in the direction of further automation, but this might mean missing out on the promise of the “right” kind of AI with better economic and social outcomes.

Keywords: automation, artificial intelligence, jobs, inequality, innovation, labor demand, productivity, tasks, technology, wages.

JEL classification: J23, J24.

---

*We thank Chris Ackerman, David Autor, Erik Brynjolfsson, Stu Feldman, Mike Piore, Jim Poterba and Hal Varian for useful comments. We gratefully acknowledge financial support from Google, Microsoft, the Sloan Foundation, the Toulouse Network on Information Technology, the Schmidt Sciences Foundation, the Smith Richardson Foundation, IBM, Accenture and the National Science Foundation.
Artificial Intelligence (AI) is one of the most promising technologies currently being developed and deployed. Broadly speaking, AI refers to the study and development of “intelligent (machine) agents”, which are machines, softwares or algorithms that act intelligently by recognizing and responding to their environment.¹ There is a lot of excitement, some hype, and a fair bit of apprehension about what AI will mean for our security, society and economy. But a critical question has been largely overlooked: are we investing in the “right” type of AI, the type with the greatest potential for raising productivity and generating broad-based prosperity? We don’t have a definitive answer right now — nobody does. But this is the right time to ask this question while we can still shape the direction of AI research and the future of work.

AI As a Technology Platform

Human (or natural) intelligence comprises several different types of mental activities. These include simple computation, data processing, pattern recognition, prediction, various types of problem solving, judgment, creativity, and communication. Early AI, pioneered in the 1950s by researchers from computer science, psychology and economics, such as Marvin Minsky, Seymour Papert, John McCarthy, Herbert Simon and Allen Newell, sought to develop machine intelligence capable of performing all of these different types of mental activities.² The goal was nothing short of creating truly intelligent machines. Herbert Simon and Allen Newell, for example, claimed in 1958 “there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which the human mind has been applied.”³

These ambitious goals were soon dashed. AI came back into fashion in the 1990s, but with a different and more modest ambition: to replicate and then improve upon human intelligence in pattern recognition and prediction (pre-AI computers were already better than humans in computation and data processing). Many decision problems and activities we routinely engage in can be viewed as examples of pattern recognition and prediction. These include recognizing faces (from visual data), recognizing speech (from auditory data), recognizing abstract patterns in data we are presented with, and making decisions on the basis of past experience and current information. Though there are researchers working on “Artificial General Intelligence”, much of the research and almost all commercial applications of AI are in these more modest domains referred to as “Narrow AI” — even if the relevant applications

---

²See Nilsson (2009) for the history of AI.
are numerous and varied. The big breakthroughs and the renewed excitement in AI are coming from advances in hardware and algorithms that enable the processing and analysis of vast amounts of unstructured data (for example, speech data that cannot be represented in the usual structured ways, such as in simple, Excel-like databases). Central to this renaissance of AI have been methods of machine learning (which are the statistical techniques that enable computers and algorithms to learn, predict and perform tasks from large amounts of data without being explicitly programmed) and what is called “deep learning” (algorithms that use multi-layered programs, such as neural nets, for improved machine learning, statistical inference and optimization).

Even if we focus on its narrow version, AI should be thought of as a technology platform — there are many ways AI technology can be developed as a commercial or production technology, with widely varying applications. To some degree, this is true of all clusters of technologies, but it is more emphatically so for AI. To see this, contrast it with a related but distinct new technology, robotics. Robotics often makes use of AI and other digital technologies for processing data, but is distinguished from other digital technologies by its focus on interacting with the physical world (moving around, transforming, rearranging or joining objects). Industrial robots are already widespread in many manufacturing industries and in some retail and wholesale establishments. But their economic use is quite specific, and centers on automation of narrow tasks, that is, substituting machines for certain specific activities and functions previously performed by humans.

Implications of Technology for Work and Labor

How do new technologies impact the nature of production and work? Employment and wages of different types of workers? The standard approach, both in popular discussions and academic writings, presumes that any advance that increases productivity (value added per worker) also tends to raise the demand for labor, and thus employment and wages. Of course, technological progress might benefit workers with different skills unequally and productivity improvements in one sector may lead to job loss in that sector. But even when there are sectoral job losses, the standard narrative goes, other sectors will expand and contribute to overall employment and wage growth.

This view is critically underpinned by the way in which the economic impact of new

---

4Bessen et al. (2018) report that many commercial AI startups view their technology as capable of “enhancing human capabilities”, while many others recognize that their technologies have a significant automation component.

technology is conceptualized — as enabling labor to become more productive in pretty much all of the activities and tasks that it performs. Yet, this not only lacks descriptive realism (what technology makes labor uniformly more productive in everything?), but may paint an excessively rosy picture of the implications of new technologies. Indeed, in such a world Luddites’ concerns about the disruptive and job displacing implications of technology would be misplaced, and they would have smashed all of those machines in vain.

The reality of technological change is rather different. Many new technologies — those we call *automation technologies* — do not increase labor’s productivity, but are explicitly aimed at replacing it by substituting cheaper capital (machines) in a range of tasks performed by humans. As a result, automation technologies always reduce the labor’s share in value added (because they increase productivity by more than wages and employment). They may also reduce overall labor demand because they displace workers from the tasks they were previously performing. There are countervailing effects to be sure: some of the productivity gains may translate into greater demand for labor in non-automated tasks and in other sectors. But even with these effects factored in, automation always reduces the labor share in value added.

This discussion clarifies an important point: in an age of rapid automation, labor’s relative standing will deteriorate and workers will be particularly badly affected if new technologies are not raising productivity sufficiently — if these new technologies are not great but just “so-so” (just good enough to be adopted but not so much more productive than the labor they are replacing). With so-so automation technologies, labor demand declines: the displacement is there, while powerful productivity gains contributing to labor demand are missing.

Is this far-fetched? Not really. We have previously studied the implications of one of the most important automation technologies, industrial robots. Industrial robots are not technologies aimed at increasing labor’s productivity but are designed to automate tasks that were previously performed by production workers on the factory floor. The evidence is fairly clear that industries where more industrial robots are introduced experience declines in labor demand (especially for production workers) and sizable falls in their labor share. More importantly, local labor markets more exposed to industrial robots, such as Detroit MI or Defiance OH, have significantly lower employment and wage growth. Furthermore, the declines in wages and employment fell much more heavily on workers from the lower half of the earnings distribution and those with less than a college degree, thus exacerbating inequality. All of this is despite the fact that industry-level data also suggest productivity

---

6This approach is developed in Zeira (1998), Autor, Levy and Murnane (2003), Acemoglu and Autor (2011) and Acemoglu and Restrepo (2018a,b, 2019).

7Acemoglu and Restrepo (2018b).
gains from robots.

Automation is not a recent phenomenon. Many important breakthroughs in the history of technology have been centered around automation. Most notably, the spectacular advances in the early stages of the Industrial Revolution in Britain were aimed at automating weaving and spinning, and the focus then shifted to the factory floors of other industries.\(^8\) Other prominent examples of automation are the mechanization of agriculture and the interchangeable parts system of American manufacturing.

But if automation tends to reduce the labor share and has mixed effects on labor demand, why did the labor share remain roughly constant and productivity growth go hand-in-hand with commensurate wage growth over the last two centuries? To understand this relationship, we need to recognize different types of technological advances contributing to productivity growth. Historically, as automation technologies were being introduced, other technological advances simultaneously enabled the creation of new tasks in which labor had a competitive advantage. This generated new activities for labor — tasks in which human labor could be reinstated into the production process — and robustly contributed to productivity growth as new tasks improved the division of labor.\(^9\) The episode of agricultural mechanization, which started in the second half of the 19th century, vividly illustrates this pattern. Though mechanization reduced the labor share and employment in agriculture, overall labor demand rose because a range of new tasks were introduced in both manufacturing and services. In fact, this period witnessed not only the rise of clerical occupations but a range of more specialized blue-collar and white-collar jobs that increased productivity, the demand for labor and the labor share in manufacturing and services.\(^10\) Occupations featuring new tasks have been at the forefront of employment growth in the US economy in the postwar era as well.\(^11\)

This perspective then suggests a different reinterpretation of the history of technology and a different way of thinking about the future of work — as a race between automation and new, labor-intensive tasks. Labor demand has not increased steadily over the last two centuries because of technologies that have made labor more productive in everything. Rather, many new technologies have sought to eliminate labor from tasks in which it previously specialized. All the same, labor has benefited from advances in technology, because other technologies have simultaneously enabled the introduction of new labor-intensive tasks. These new tasks have done more than just reinstate labor as a central input into the production process; they have also played a vital role in productivity growth.

\(^8\)See Mantoux (1927) and Mokyr (1989).
\(^9\)Acemoglu and Restrepo (2018a).
\(^11\)Acemoglu and Restrepo (2018a).
Viewed from this perspective, employment and wage growth have been disappointing over the last two decades partly because productivity growth has been weak, and even more importantly because new tasks have failed to materialize.\textsuperscript{12} The future of work will be much brighter if we can mobilize more of the technologies that increase labor demand and ensure vigorous productivity growth.

**Varieties of AI**

This perspective provides a new way of thinking about the economic opportunities and challenges posed by AI. Most AI researchers and economists studying its consequences view it as a way of automating yet more tasks. No doubt, AI has this capability, and most of its applications to date have been of this mold — e.g., image recognition, speech recognition, translation, accounting, recommendation systems, and customer support. But we do not need to accept that this as the primary way that AI can be and indeed ought to be used.

First, if all we do is continue down the path of automation, with no counterbalancing innovations to generate new tasks, the implications for labor are depressing. It will not be the end of work anytime soon,\textsuperscript{13} but the trend towards lower labor share and anemic growth in labor demand will continue — with potentially disastrous consequences for income inequality and social cohesion.

Second, as we go deeper and deeper into AI-based automation, we are moving into areas in which human labor is quite good (for example think of image and speech recognition or hand-eye coordination), and machine productivity, at least to start with, is not always impressive, to say the least. Automation technologies aimed at substituting machines for humans in these tasks are thus likely to be of the so-so kind. As a result, we cannot even count on powerful productivity gains to increase our living standards and contribute to labor demand.

But it doesn’t have to be this way. Since AI is not just a narrow set of technologies with specific, pre-determined applications and functionalities but a technology platform, it can be deployed for much more than automation; it can be used to restructure the production process in a way that creates many new, high-productivity tasks for labor. If this type of “reinstating AI” is a possibility, there would be potentially large societal gains both in terms of improved productivity and greater labor demand (which will not only create more inclusive growth but also avoid the social problems created by joblessness and wage declines).

Consider a few examples of how AI applications can create new tasks for labor.

\textsuperscript{12}Acemoglu and Restrepo (2019).
\textsuperscript{13}See Dreyfus (1992) and Autor (2015).
• **Education**: Education is one of the areas with the least AI penetration. That may be partly because automation is not an attractive or even feasible option for most of the core tasks in education. But using AI to create new tasks would be a different way of productively deploying this new technology platform. Consider, for example, classroom teaching. This has not changed for over 200 years. A teacher teaches to the whole class, even if he or she or an aide may occasionally engage in one-on-one instruction or provide help for some subset of students. There is evidence, however, suggesting that many students have different “learning styles”, and what works for one student may not work for another, and even what works for one student in one subject will not work for him or her in every subject.14

At the moment, individualized teaching, targeted and adapted to each student or for small subsets of students, is impossible, and not just because the resources in terms of teacher time and skill are lacking. It is mostly because nobody has the information (and cannot easily acquire and process the information) to determine a student’s optimal learning style in a specific subject or topic. AI can change this. AI software can be designed to collect and process in real time data about the specific reactions, difficulties and successes students have in different subject areas, especially when taught in different styles, and then make recommendations for improved individualized teaching. The potential improvements in terms of educational productivity could be quite large (we just don’t know). Societal benefits could exceed these direct benefits as AI-powered teaching methods may do better in terms of providing students with skills that will be more valued in future labor markets (rather than the more backward-looking curricula and teaching emphasis currently prevailing in schools). Developing and deploying such technologies would also increase the demand for human labor in teaching — we would need more teachers with diverse skills to do the individualized teaching, even with help from AI software and other technologies.

• **Healthcare**: The situation in healthcare is similar. Though there has been more effort to introduce digital technologies into healthcare, the focus has not been on creating tasks in which humans can be productively employed (in fact, some of the uses of AI, for example in radiology, are very much in the mold of automation). AI applications that collect and analyze information can significantly empower nurses, technicians and other healthcare providers to offer a wider range of services and more real-time health advice, diagnosis and treatment. The benefits in terms of greater labor demand and productivity are very similar to the education case.

• **Augmented reality:** The third area in which the use of AI can significantly change the production process in a way that may be favorable to labor is through augmented and virtual reality technologies in manufacturing. Most advanced manufacturing technologies of the last three decades have focused on automation. But companies such as Amazon and Tesla have discovered that automating all factory-floor and manual tasks is not economically rational, because some tasks are still more productively performed by humans. One difficulty facing companies introducing industrial robots, however, is that these new technologies do not necessarily integrate well with humans for at least two reasons. First, most robotics technology is cordoned off from workers because of safety concerns. Second, human work may not mesh with the degree of precision required by robotics technology. Augmented reality technologies — which use interactive interfaces in order to increase the ability of humans to perceive, monitor and control objects — might enable workers to work alongside machines and perform high precision production and integrated design tasks. This will not just help workers keep some of the tasks that might have otherwise been automated; it could also create new tasks in which humans, augmented by digital technology and sensors, can be employed and contribute to productivity.\(^{15}\)

Notably, the examples of new tasks mentioned above go well beyond what are sometimes emphasized as “enablers” of AI — human tasks involved in training and monitoring new machines as they automate what the rest of us do. This is critical; work in just enabling AI is unlikely to generate sufficient new tasks and demand for human labor to undergird broad-based prosperity.

**Why the Wrong Kind of AI?**

If there are potentially productive and profitable uses of AI beyond simple automation, can we count on market forces and innovation by existing companies to take us there? Is there any reason to worry that AI applications with the promise of reinstating human labor will not be exploited and resources will continue to pour instead into the wrong kind of AI?

Economists tend to place great trust in the market’s ability to allocate resources in the most efficient way. But most experts recognize that the market’s star doesn’t shine as brightly when it comes to innovation. There are several reasons for market failures in innovation in general, as well as some specific reasons that are important in the context of AI.

\(^{15}\)See Ong and Nee (2013) and https://www.ge.com/reports/game-augmented-reality-helping-factory-workers-become-productive/).
Innovation creates externalities — not just the innovator, but the workers that use the new technology, the firms that adopt it, and most importantly other firms and researchers building on it in the future will benefit from it. Markets do not do a good job in the presence of such externalities.

Markets struggle when there are alternative, competing technological paradigms. When one paradigm is ahead of the other, both researchers and companies tend to follow that paradigm, even if an alternative could be more productive. Moreover, in such a situation, once the wrong paradigm pulls ahead, it may be very difficult to reverse this trend and benefit from the possibilities offered by the alternative paradigm. To the extent that different approaches to AI constitute alternative, competing paradigms, our trust in the market mechanism getting it right should be even lower.\textsuperscript{16}

To correct market failures in innovation, the US government has historically used public-private partnerships to encourage socially beneficial research. It has played an important role in many leading technologies, including the Internet, sensors, pharmaceuticals, biotech and nanotechnology.\textsuperscript{17} But more recently, the US government has been more frugal in its support for research and more timid in its determination to steer the direction of technological change. Part of this shift is due to the reduction in resources devoted to government support of innovation and the increasingly dominant role of the private sector in setting the agenda in high-tech areas (can government officials and researchers meaningfully influence the direction of inventive activity in Silicon Valley?). This shift will further discourage research related to future promise (that is not immediately reflected in profitability) and other social objectives (such as employment creation).

Innovation does not just respond to economic incentives. Several noneconomic rewards affect what types of technologies attract the attention and imagination of researchers. It is possible that the ecosystem around the most creative clusters in the United States, such as Silicon Valley, excessively rewards automation and pays insufficient attention to other uses of frontier technologies. This may be partly because of the values and interests of leading researchers (consider for example the ethos of companies like Tesla that have ceaselessly tried to automate everything). It is also partly because the prevailing business model and vision of the large tech companies, which are the source of most of the resources going into AI, have focused on automation and removing the (fallible)\textsuperscript{16}See Nelson and Winter (1977), Dosi (1982) and Acemoglu (2012).\textsuperscript{17}Mazzucato (2015).
human element from the production process. This last consideration may have become even more critical as the vast resources of several leading companies are pouring into academia and shaping the teaching and research missions of leading universities. It is then no surprise that the best minds in the current generation are gravitating towards computer science, AI and machine learning, but with a heavy emphasis on automation. An ecosystem that is biased would become much more stifling for the direction of technological change when it becomes all-encompassing.

- There are also additional factors that may have distorted choices over what types of AI applications to develop. The first one is that if employment creation has a social value beyond what is in the GDP statistics (for example, because employed people are happier and become better citizens, or because faster growth in labor demand improves income inequality), this social value will be ignored by the market. The second is related to the tax policies adopted in the United States and other Western nations, which subsidize capital and investment while taxing employment. This makes using machines instead of labor more profitable, and these profits encourage not just automation but also automation research. Finally, and complementing these factors, to the extent that firms take into account the cost of labor (the wage rate), which tends to be higher than the social opportunity cost of labor because of imperfections in the labor market, they will have additional incentives for adopting and developing automation technologies beyond what is socially optimal.\(^\text{18}\)

- Another set of factors blocking the path of novel AI applications reinstating labor is that these new technologies might need critical complementary inputs that are not forthcoming. Take the example of education mentioned above. It is not only that developing AI to create new labor-intensive tasks in education is not viewed the frontier or one of the “cool” areas of research, say compared to facial recognition. It is also that complementary skills and resources to make this type of reinstating AI profitable may be missing completely. Educational applications of AI would necessitate new, more flexible skills from teachers (beyond what is available and what is being invested in now), and they would need additional resources to hire more teachers to work with these new AI technologies (after all, that is the point of the new technology, to create new tasks and additional demand for teachers). In the case of healthcare, limited resources are not the problem (the share of national income devoted to health is continuing to grow), but the requisite complementary changes are likely to be organizational. In fact,

\(^{18}\)Acemoglu and Restrepo (2018a).
highlighting other barriers to the use of new technologies to create new tasks, the way that hospitals, insurance companies and the whole medical profession, as represented by the American Medical Association, is organized is likely to be in the way. If empowering, and increasing the productivity of, nurses and technicians is perceived to reduce the demand for the services of doctors or challenge the current business model of hospitals, it will be strenuously resisted.

All in all, even though we currently lack definitive evidence that research and corporate resources today are being directed towards the “wrong” kind of AI, the market for innovation gives no compelling reason to expect an efficient balance between different types of AI. If at this critical juncture insufficient attention is devoted to inventing and creating demand for, rather than just replacing, labor, that would be the “wrong” kind of AI from the social and economic point of view. Rather than undergirding productivity growth, employment and shared prosperity, rampant automation would contribute to anemic growth and inequality.

Conclusion

Artificial Intelligence is set to influence every aspect of our lives, not least the way production is organized in modern economies. But we shouldn’t assume that, left to its own devices, the right types of AI will be developed and implemented. Though many today worry about the security risks and other unforeseen (often non-economic) consequences of AI, we have argued that there are prima facie reasons for worrying about the wrong kind of AI from an economic point of view becoming all the rage and the basis of future technological development. The considerable promise of AI implies that we need to devote care and serious thought to its implications and to the question of how best to develop this promising technology platform — before it is too late.
References


