Preparing the Workforce for an Uncertain AI Future

Dr. Anton Korinek
Professor of Economics, University of Virginia and Darden School of Business
Economics of AI Lead, Centre for the Governance of AI

Written Statement to the U.S. Senate AI Insight Forum on Workforce
Washington, D.C.
November 1, 2023

Majority Leader Schumer, Senators Rounds, Heinrich, and Young, and Distinguished Members of the Senate, thank you for organizing this important event today and for inviting me to it. I particularly appreciate the opportunity to speak about the workforce implications of advanced AI on a day when much of the attention of the world's AI community is focused on the AI Safety Summit in the UK. Ensuring the safety of AI systems is important, but not enough. We also need to prepare for the workforce implications of advanced AI, which I believe will be one of the greatest challenges for our political system in the coming years (Bell and Korinek, 2023).

In the following,

- I will lay out why advances in AI are likely to usher in a new era of cognitive automation,
- I will emphasize the breakneck speed of advances in frontier AI systems and the resulting uncertainty about the workforce implications of future AI,
- I will observe that progress creates the possibility of greater shared prosperity, but that getting there requires policy action, both by steering AI progress and by updating the way our economic institutions allocate income,
- I will propose that the uncertainty calls for scenario planning, including transformative scenarios in which AI can automate most work within 5 to 20 years,
- and I will end with two recommendations for the U.S. Senate to prepare for the workforce implications of future advances in AI.

We are entering a new era of cognitive automation. For much of the past half-century, the U.S. economy experienced skill-biased technological change that favored workers with higher education while leaving less-educated workers behind (Goldin and Katz, 2010). As a result, economists advised that the best preparation for our workforce was to pursue additional education. Americans followed suit, and those in the workforce with higher education reaped the rewards.

All of this may now be changing. The latest generation of AI systems, called Generative AI, are "machines of mind" whose general-purpose nature differs starkly from earlier technological advances. Whereas 20th-century innovations like assembly lines or factory robots automated physical tasks, we are now automating increasingly complex cognitive work. Ironically, the workers most affected are those who pursued more education (Eloundou et al., 2023; Felten et al., 2023). For instance, computer programmers are among the most impacted categories of workers (Peng et al., 2023).

AI systems are advancing at a breakneck speed. Over the past 18 months, AI systems have acquired many new capabilities that will have profound implications for the workforce. They have suddenly crossed the threshold where they have become useful for a wide range of cognitive tasks (Dell'Acqua et al., 2023). Perhaps it is no coincidence that this is occurring roughly at the time when the computational complexity of the leading AI systems is starting to approach estimates of the complexity of human brains (Carlsmith, 2020). The recent pace of progress has astonished even some of the field's pioneers, such as Turing Award winners Yoshua Bengio and Geoffrey Hinton (De Vynck, 2023, Knight, 2023).

The rapid rise of AI is the result of more than a decade of exponential growth in the amount of computing power devoted to training the leading AI systems, which has been quadrupling every (Epoch, 2023), and by rapid progress in the algorithms used to train these systems (Hernandez and Brown, 2020, Erdil and Besiroglu, 2022). Together, these advances imply that the *effective* computing power used to train the most advanced AI systems has been growing by an estimated *factor of ten per year* over the past decade.

Crucially, there is no end in sight for this rapid growth. The described trends are highly likely to continue for at least several more years. Nobody knows what this will imply for the capabilities of the leading AI systems.

We need to steer technological progress to deliver shared prosperity. Economists define technological progress as the ability to obtain greater economic output for a given set of inputs (Korinek and Stiglitz, 2019). If this greater output is distributed equitably, everyone could benefit. And if the early predictions about the productivity effects of recent advances in AI prove correct, then everyone could benefit quite a lot.

Yet while how much we can produce is determined by technology, how it is distributed is shaped by our institutions. Even if progress makes it possible, shared prosperity does not automatically materialize without thoughtful policy guidance (Klinova and Korinek, 2021). Historically, technological leaps often created winners and losers, with entire groups left behind. For example, over the past four decades, blue-collar workers in the American heartland have experienced what it means to be left behind by technological progress, no matter how hard they work. Cognitive workers may be next in line.

When the potential for greater shared prosperity is not realized, it represents a failure of our economic institutions. It means that we have wasted the opportunities that technological progress presented to uplift all Americans. With automation accelerating, policies that promote shared prosperity may be not only a moral obligation but also a necessity to safeguard the survival of our democracy. But how to best do so depends on the shape of future progress in AI.

The speed and ultimate trajectory of how AI will automate jobs remain uncertain. Right now, we are in the "honeymoon period" of general-purpose AI. Cognitive workers can derive significant productivity gains from the latest AI systems (as I have explored, for example, in the area of economics; see Korinek, 2023). Few have suffered job disruption so far. Businesses are still working to figure out the optimal deployment of the current generation of AI. This honeymoon period may last for some time and may give rise to a beneficial period of inclusive growth. However, leading tech companies are releasing new AI systems with more powerful capabilities

every month, and the balance could swing abruptly from enhancing workers with technology to replacing them. Moreover, AI systems are also making robots more capable, which may soon lead to further automation of blue-collar jobs.

Observing the pace of advances, a growing number of AI experts are predicting that AI may soon be able to perform *all* cognitive work, making human workers obsolete. AI that can match all cognitive capabilities of humans is often referred to as Artificial General Intelligence (AGI). After quitting his job at Google in April 2023, AI pioneer Geoffrey Hinton declared: "I have suddenly switched my views on whether these things are going to be more intelligent than us" (Heaven, 2023). He anticipates AI to become smarter than humans in 5 to 20 years.

If AI systems can perform all cognitive work and are cheaper, market wages of cognitive workers would tend to fall to the level of AI systems' operating cost. And if AI systems in combination with advanced robots can perform virtually all human jobs in the future, wages for all human labor would tend to decline significantly (Korinek and Juelfs, 2022). A certain level of demand for human jobs would likely remain from two sources: because consumers may have an intrinsic preference for humans to perform certain services (e.g., law-making) and because our society may choose to retain certain jobs with humans for ethical reasons or for reasons of better controlling AI systems. Yet it is questionable whether these two categories of jobs would generate sufficient demand to maintain the current level of wages.

Other experts - including some economists who have studied earlier rounds of innovations - remain deeply skeptical based on a different perspective of how powerful AI might become in comparison to human brains. They point out that such a development would contradict our experience of the past two hundred years. They observe that in the past, the economy has always created new jobs when old jobs were automated. We have indeed regularly invented new jobs in the past, leaving automated tasks to machines and leveraging our unique human abilities in areas that were hitherto unautomated. However, at a fundamental level, all new jobs consist of recombinations of existing human abilities. Once machines can acquire all those abilities, they would be as good as humans at any new jobs too.

We need scenario planning to prepare for the future. The starkly divergent perspectives on AI's future trajectory reflect how much uncertainty we as a society currently face. Given this uncertainty, the prudent approach is to avoid putting all eggs into one basket and betting on any single scenario. Instead, we need to prepare for a range of possibilities through robust scenario planning. The following three scenarios span a wide range of the possible trajectories of how AI may affect the workforce:

I) The historical scenario: if AI continues to evolve along the same lines as earlier general-purpose technologies such as the steam engine or electricity, then it will automate a range of work tasks but also create new opportunities for the affected workers. It will increase productivity, and it will have the potential to uplift the workforce.

II) The baseline scenario - AGI in 20 years: if AI continues to steadily advance to reach human-level abilities across all work tasks 20 years from now, then it will fundamentally disrupt the workforce. Yet it would also allow for economic growth to be turbo-charged as machines can produce economic output without the need for human inputs.

III) The aggressive scenario - AGI within 5 years: if advances in AI accelerate even further in coming years and all cognitive work can be performed by machines within 5 years from now, the disruption to our workforce would be starkest. In this scenario, it is likely that white-collar work will be displaced significantly faster than blue-collar work as robotics would need to catch up to the cognitive capabilities of AI.

I currently assign a probability of at least 10 percent to each of these scenarios, although I am highly uncertain.

We need a plan for economic policy that works across scenarios. In the historical scenario, many of the workforce policies that have proven important in recent decades will continue to apply. Educating our workforce - including education on how to best leverage the latest generation of AI systems - will allow them to be more productive. Regulating new harms from AI - for example, excessive surveillance and dehumanizing working conditions - will be of utmost importance. The cognitive nature of AI would likely increase the relevance of trade skills vs cognitive skills. This may lead to a healthy rebalancing after several decades in which cognitive workers benefited more from growth.

The AGI scenarios, by contrast, would pose grave challenges for our workforce and our policy institutions. Our society is currently organized around work as the primary way of earning income, and AGI would threaten to undermine work. Without fundamental reforms to our economic institutions, AGI would likely deliver mass impoverishment, with all the resulting implications for our political system. Useful reforms may include elements of universal basic incomes (UBI), universal basic capital (UBC), or other mechanisms that allow all Americans to participate in the prosperity generated by advanced AI.

Crucially, however, such mechanisms must be designed in a manner that works under all possible scenarios. For example, a significant UBI is not very desirable in our current system and would not be desirable in the historical scenario since the existing social safety net offers better targeting at a lower cost.

What I advocate instead is a "seed UBI" that only provides a nominal income for now but ensures that the infrastructure for providing universal payments is put in place. The program can be designed such that it automatically scales up if one of the transformative scenarios materializes and the need arises. Creating this infrastructure takes significant time, and given the potential for rapid advances in AI that may create the need for it, now is the time to act on it. An additional benefit of this infrastructure is that it could be used for payments in the event of other national emergencies, such as economic crises, natural disasters, or national security events.

Although the challenges that AGI would pose for the workforce would be momentous, there are also reasons for optimism: First, we should remember that work is a means to an end, not an end in itself. If technology could provide a decent standard of living without obligatory work, it could allow Americans to find new sources of meaning and live more fulfilled lives with deeper relationships. Second, AGI would enable the economy to produce vastly greater wealth with less human labor (Trammell and Korinek, 2023). If even just a fraction of this wealth accrues to displaced workers, AGI *could* make everyone better off by expanding the economy even as it shrinks the portion going to human work.

Recommendations: Given the unprecedented uncertainty surrounding AI's future direction, I recommend the following measures to optimally prepare and assist the American workforce:

1. **Establish a commission to examine how different AI scenarios would affect the workforce.** The commission should encompass both AI experts who lay out scenarios - including different AGI scenarios - and economists who evaluate the economic implications. First, the commission should evaluate how different scenarios for advances in AI will affect the US workforce and by extension income distribution. Second, the commission should evaluate how well the existing social safety net would protect the American workforce under those scenarios. Third, it should examine how to reform the social safety net and tax code to be prepared for transformative AI scenarios, including reforms to unemployment insurance, EITC, Social Security, Medicare, Medicaid, SNAP, and TANF. Fourth, it should analyze how to raise the requisite tax revenue if the role of labor declines under such scenarios.

2. **Plan for operationalizing reforms to the social safety net and tax code:** Instruct the U.S. Department of Treasury and Department of Health and Human Services to develop an operational plan within 12 months for how to implement reforms to the safety net and tax code that may be needed in an AGI scenario. The plan should delineate the legislative and regulatory actions needed to implement such reforms, including:
   - Defining eligibility criteria
   - Designing distribution mechanisms leveraging existing infrastructure like Social Security
   - Identifying funding sources
   - Incorporating adjustment mechanisms that scale benefits based on economic developments, productivity growth, and the pace of automation

   This operational plan will provide Congress with the necessary groundwork to make informed decisions. I believe putting these measures in place proactively is imperative to ensure that technological progress benefits all Americans.

As many of the experts in the first two AI Insight Forums have emphasized, it is crucial to align AI systems with human goals. A critical dimension of this alignment is that AI does not undermine the livelihood of Americans. AI that degrades and impoverishes us without generating alternative sources of income and wealth would be a significant failure mode. Yet there is also much promise that with the right policies, AI could make all Americans much better off and deliver greater shared prosperity. This is a future I dream of for my children. With careful governance, technological progress may make this dream a reality.