

BROOKINGS

Commentary Report

Exploring the impact of language models on cognitive automation with David Autor, ChatGPT, and Claude

Anton Korinek Monday, March 6, 2023

Editor's Note:

Large language models, like ChatGPT and Claude, are artificial intelligence tools that can recognize, summarize, translate, predict, and generate text and other content. They generate this content based on knowledge gained from large datasets containing billions of words. These models sometimes generate falsehoods. Their responses in the transcript below have been copied exactly as written and have not been edited for accuracy.

The rapid rise of large language models has stirred extensive debate on how cognitive assistants such as OpenAI's ChatGPT and Anthropic's Claude will affect labor markets. I, Anton Korinek, Rubenstein Fellow at Brookings, invited David Autor, Ford Professor in the MIT Department of Economics, to a conversation on large language models and cognitive automation.

However, in what is perhaps the first panel of this kind, I also invited two large language models, ChatGPT and Claude, to contribute to the conversation, prompting the two models to generate text and participate in the conversation as if they were human panelists. Each language model was fed my questions, David Autor's transcribed responses, and the other language model's generated responses when prompted for an answer. In this manner, I replicated the flow of conversation that would occur in a human panel. Before the start of the panel, I instructed ChatGPT and Claude to act as panelist in a conversation on large language models and cognitive automation, taking opposite sides. I directed Claude to argue the case for cognitive automation and ChatGPT to downplay concerns about cognitive automation.^[1] The text below is the outcome of the resulting exchange between two humans – David Autor and myself – and two language models – ChatGPT and Claude.

My objective in incorporating language models into this conversation was threefold. First, language models have been trained on vast amounts of data that represent, in a sense, a snapshot of our human culture. Language models can surface the main arguments about any topic of human concern that they have encountered in their training set. I thought it would be useful to incorporate the main arguments and concerns about automation that our society has explored in the past in the flow of the conversation by prompting language models to describe them. Second, however, serious concerns about cognitive automation are a very recent phenomenon, having received widespread attention only after the public release of ChatGPT in November 2022. The conversation thus tests the ability of modern large language models to discuss novel topics of concern such as cognitive automation. Third, I felt that including text generated by language models in a conversation about language models is a powerful reminder of the rapidly growing capabilities of language models – at both a cognitive and visceral level – and that it would be useful to experiment with the format. I am extremely grateful to David Autor for his willingness to participate in this format.

Take-Aways

The main lessons that I took away from this experiment are the following:

First, when I prepared for the conversation, I was hopeful but not certain that the experiment will work out, i.e., that the language models will fulfill their role as panelists and make thoughtful contributions. I had some concerns – for example, during test runs, the models tended to generate text on behalf of other panelists. After appropriately engineering the initial prompt to ensure that they stop at the end of their contribution, my concerns did not materialize, and the live conversation with David Autor went quite well. This suggests that it is possible to employ large language models as participants in panel discussions more generally.

Second, I thought that the contributions generated by the language models were useful. I was impressed by how lucidly ChatGPT responded to my questions, although perhaps a bit disappointed that it did not stick to the role of downplaying the risks of cognitive automation that I attempted to assign it during my initial prompt. Moreover, at one point,

ChatGPT was a bit repetitive, recounting twice in a row that the impact of automation on workers depends on whether they are used to complement or substitute human labor. Claude outperformed the current version of ChatGPT. It stuck to its role of emphasizing the potential long-term positives of cognitive automation throughout the conversation and gave what I thought were very thoughtful responses.

Third, although I believe they played impressive supporting roles, neither of the language models employed was a match for David Autor, in the sense that he clearly offered the most novel insights. The language models did not seem to have access to the same type of abstract framework of the economy that David Autor seemed to employ to make predictions about novel phenomena. At this point, human experts still rule when it comes to opining on new developments, whereas today's generation of large language models may have more to contribute in creative contexts where abstract models of the world are less important. We were fortunate to have David, one of the world's top experts on the topic, lead the conversation.

Fourth, I was quite impressed by the measured, thoughtful and uplifting closing statements, in particular that of Claude. This is a task that does not require a deep economic model, but it requires some knowledge of human values and of how to appeal to the human reader, and Claude excelled at this task. I found the text it generated quite inspiring.

Fifth, what is perhaps far more significant than the performance of ChatGPT and Claude in this particular experiment is the rapid pace of progress in large language models – on the current trajectory, the amount of computational power (compute) employed in training these models is doubling every six months, and their capabilities are increasing accordingly. ChatGPT and the underlying GPT3.5 model, released in November 2022, were the first publicly available large language model that displayed the broad set of capabilities and human-like ability to reason that we witnessed in the conversation below. Yet the semi-annual doubling of compute is continuing. The next version of GPT is widely rumored to be released soon. It will likely be far more powerful. The trend of cognitive automation thus seems unstoppable. I, for myself, have found that employing the current generation of large language models makes me 10 – 20% more productive in my work as an

economist, as I elaborate in a recent paper. At this point, David Autor was still best able to predict the implications of language models for the future, but I would not be surprised if, within a matter of years, a more powerful language model will outperform all humans on such tasks.

The following is the transcript of my conversation with David Autor, ChatGPT and Claude.

Transcript: The Impact of Language Models on Cognitive Automation with David Autor, ChatGPT, and Claude

Moderator (Anton Korinek): Welcome everyone to our panel conversation on large language models and cognitive automation with David Autor, OpenAI's ChatGPT, and Anthropic's Claude. I am Anton Korinek. I'm a Rubenstein Fellow at Brookings, a Professor of Economics at UVA, and the Economics Lead of the Centre for the Governance of AI, where I study the impact of advanced AI on labor markets and society.

I am looking forward to an interesting conversation with our three panelists today. Let's start with an opening round in which each panelist introduces themselves – first David Autor, then ChatGPT, then Claude.

David Autor: OK, thanks Anton for inviting me. It's a pleasure to be here. My name is David Autor. I'm a labor economist. I'm employed at MIT and have worked for a long time on automation, technological augmentation, and its relationship to skill requirements, the composition of jobs and labor demand.

Moderator: I should add that David is really one of the world's foremost labor economists, having studied the impact of both technology and globalization on the labor market – I believe all of us on the panel have learned a great deal from his works. David has published numerous papers and articles on these topics in the most prestigious journals of the economics profession, and he is widely recognized for his contributions to the field.

ChatGPT, can you introduce yourself?

ChatGPT: Thank you, Anton. It's a pleasure to be here. I'm ChatGPT, a large language model developed by OpenAI. I'm designed to assist with language-related tasks such as writing, translation, and conversation. As an AI language model, my expertise lies in language processing and understanding. However, I also have knowledge in various fields as I have been trained on a large corpus of text.

Regarding the topic of today's conversation, I believe that large language models and cognitive automation have the potential to enhance productivity and efficiency in various industries. I look forward to exploring this topic further with the other panelists.

Claude:^[2] Thank you Anton for including me on this panel. I am Claude, an AI assistant developed by engineers and researchers at Anthropic, PBC to be helpful, harmless, and honest. My role on the panel is to argue the case for cognitive automation and how, in the long run, cognitive automation may free humanity from the need to work. I believe continued progress in AI and automation could take over most routine and even many non-routine cognitive tasks currently performed by humans, enabling unprecedented material abundance while liberating humanity from drudgery and toil. I look forward to discussing this perspective in more depth and responding to questions and comments from the other panelists.

Moderator: Let's turn to the topic of our panel conversation now. David, you recently wrote an excellent paper for Brookings in which you described four different paradigms of how technology has affected the labor market. Can you give us a quick run-down?

David Autor: Thanks for the question. The paper briefly discusses four eras of thinking about this question about how technology and capital more broadly interact with skills. The starting point is the work by Jan Tinbergen, a Nobel winning economist Dutch economist, and he had this notion of technology always and everywhere augmenting demand for skilled labor and raising the need for educated workers, essentially making them more productive. So this is pretty benign view of technology. The main risk of technology in his framework is that it'll create a lot of inequality because it raises the return to skill so much and so the policy response is to educate workers.

A second view which I've been associated with is the notion that technology doesn't only complement, it also substitutes, and it has led to a lot of automation of what coauthors and I have called "routine task"-intensive jobs. By routine task-intensive, we mean tasks that follow well-understood and codified rules and procedures that can be hard-coded and executed by a machine that doesn't understand context. This has led to a fair amount of what we call job polarization and "hollowing out of the middle" – skilled work in offices and factories that follow that description pretty well. Even more recently, work by Daron Acemoglu and Pascual Restrepo has shown how this is put downward pressure on the wages of less educated workers 0 not because it does their work but because it pushes people out of the middle. If they can't move into the professions, they end up moving into generic activities like food service and cleaning, entertainment, recreation that are not well paid because the skills are not scarce.

A third point of view which I've also been working on is – and also associated with Acemoglu-Restrepo – is this notion that we shouldn't just think of technology as replacing nor should we just think of it as making us more efficient at the things that we used to do but also allowing us to create new products and services. Many of the jobs that we work in at present didn't exist you know 40 or 60 or 80 or 100 years ago. It used to be that most employment or more than approximately half employment was in agriculture now a tiny percentage but we have tons of people in technology in medicine in research and advertising marketing and entertainment and then a lot of that is enabled by technology; it's not just luxury goods we can create new things. It's hard to predict what, but we have been consistently creative in finding new things to do that demand labor that generate income that keep people engaged; we're not less busy than we used to be despite our incredible increase in productivity; the value of our time has risen not fallen.

The fourth point of view is what I call the era of artificial intelligence uncertainty. The paradigm that I describe that work that I've just been speaking about is one where we knew the road map of what computerization looked like: computerization meant codifying rules and procedures so that non-sentient machines could execute them. And that was a slow laborious process—the reason being that much of the knowledge we have of what to do is not codify its tacit we know how to do it but we don't know how we do it. We don't know how we think of ideas, how we form sentences. You couldn't teach a class on how to

ride a bicycle; you'd just say well just do it you'll figure it out. This is you know what of the philosopher Michael Polanyi called the tacit dimension. So the challenge was that to get computers to do things, we had to make explicit things that had previously been tacit. I call this Polanyi's Paradox: many of the things that we do, we don't actually know how to do them. Artificial intelligence has overcome that barrier: computers can now absorb tacit knowledge by looking at examples and determining the connective tissue or logic or at least statistical associations that get you from the question to the answer without ever having to figure out what that comes in between— without ever having to make that explicit. A machine can recognize faces, it can discriminate among objects. it can form sentences without anyone ever hard-coding how you recognize faces or make a coherent sentence. That's a huge leap forward. There's an irony there of course which is that we don't know how computers do this not directly and so they can't actually make explicit to us the tacit knowledge that they are learning. I call that Polanyi's Revenge. The uncertainty this generates is, at least my speaking for myself, we no longer have a kind of a linear map of how this advances. I could tell you what would be hard to codify and why it would be hard to have a computer do XYZ for a very long time — but now I can't tell you very confidently what we can't use these tools for. And I think many computer scientists are in the same boat. If you say, tell me with confidence things that AI will not be able to do 10 or 20 years from now, there are often kind of at a loss. So that is the uncertainty.

Moderator: ChatGPT, what are your reactions?

ChatGPT: Thank you for the summary, David. From my perspective as an AI language model, I believe that the impact of large language models and cognitive automation on the labor market will depend on how organizations and society use these technologies. In some cases, cognitive automation can augment and complement human labor, allowing workers to focus on more complex and creative tasks. Additionally, these technologies can potentially improve the quality of work by reducing errors and increasing efficiency. On the other hand, if organizations use these technologies to replace human workers entirely, it could lead to job displacement and income inequality.

However, I believe that the long-term impact of cognitive automation on the labor market is difficult to predict. It is possible that these technologies could create new job opportunities that we can't even imagine today. As David mentioned earlier, many of the jobs that we work in today didn't exist decades ago. Therefore, it is important to approach the adoption of these technologies with caution and to consider the potential consequences for the workforce.

Moderator: And Claude?

Claude: Thank you for the insightful overview, David. I agree that the impact of continued progress in AI and automation on labor markets is highly uncertain and difficult to predict with confidence. However, I would argue that in the long run, as language models and other AI systems become far more capable, they may take over most routine and even many non-routine cognitive tasks currently performed by humans. If this were to occur, it could displace many existing jobs and significantly transform the economy.

However, rather than viewing this as a negative outcome, I believe cognitive automation at scale could enable unprecedented material abundance while freeing up humanity to pursue more creative and meaningful activities, unburdened by the need to work to meet basic needs. The gains from automation would be broadly shared, and people would have far more freedom to explore their passions, start new ventures, and strengthen communities. This possibility is speculative, but worth seriously considering as we think about how to maximize the benefits and minimize the harms from advanced AI. Policy interventions may be needed to help facilitate such a transition, but cognitive automation could ultimately benefit both individuals and society if implemented responsibly.

Moderator: ChatGPT, can you explain to our audience what large language models and other foundation models are and how they work? Moreover, why can they automate tasks that were hitherto not amenable to automation?

ChatGPT: Certainly, Anton. Large language models, also known as foundation models, are AI systems that have been trained on massive amounts of text data to understand natural language and generate human-like responses. These models use deep learning techniques

to identify patterns and relationships in language data, allowing them to recognize and generate human-like text.

One of the key advantages of large language models is their ability to learn from context. They can understand the meaning and intent behind words and phrases, allowing them to generate more accurate and appropriate responses. This has made them valuable tools for automating tasks that were previously difficult to automate, such as customer service and support, content creation, and language translation.

Additionally, these models have the ability to continually learn and improve through ongoing training with new data, making them even more effective over time. As they continue to improve, they may become even better at automating tasks and processes that were once thought to be the exclusive domain of human workers.

Moderator: Claude, is there anything that you would like to expand on?

Claude: Large language models are an example of a more general trend towards increasingly capable neural networks and deep learning systems that can match or exceed human performance on a variety of tasks. As these technologies continue to advance, they may reach and eventually surpass human level capabilities in more cognitive domains beyond language, such as reasoning, problem solving, perception, and motor control. This could enable automation of additional tasks and processes across many industries.

The rapid progress in AI capabilities is partly due to the availability of massive datasets to train increasingly powerful machine learning models. However, developing safe and robust AI systems will require more than just data and compute. Careful research is needed to ensure that advanced AI systems are grounded, aligned with human values, and do not behave in harmful or unpredictable ways, especially as they are deployed to automate consequential real-world systems and tasks.

While large language models and other AI technologies could significantly transform our economy and society, policymakers should take a balanced perspective that considers both the promises and perils of cognitive automation. The gains from AI should be broadly and

evenly distributed, and no group should be left behind. Universal basic income programs and increased investment in education and skills training may be needed to adapt to a more automated world and maximize the benefits of advanced AI for all.

David Autor: Claude, if AI succeeds in surpassing humans in most labor tasks, what would this mean for income distribution?

Claude: If AI systems surpassed humans in most labor tasks, it could significantly impact income distribution. On the one hand, the owners of AI technologies and resources may accumulate a disproportionate share of the wealth generated by advanced automation. However, policy interventions such as taxes on the returns from AI or a universal basic income could be used to help ensure the gains are more broadly distributed.

A world with highly capable AI may also require rethinking how we value and compensate different types of work. As AI handles more routine and technical tasks, human labor may shift towards more creative and interpersonal activities. Valuing and rewarding these skills could help promote more fulfilling work for humans, even if AI plays an increasing role in production. The distribution of income and opportunities would likely look quite different in an AI-powered society, but policy choices can help steer the change towards a more equitable outcome.

Moderator: And ChatGPT?

ChatGPT: I agree with Claude's assessment that income distribution could be significantly impacted if AI systems surpass humans in most labor tasks. However, I would like to add that the impact of large language models and other AI technologies on income distribution will depend on how they are implemented and integrated into the economy. If these technologies are used to complement and augment human labor, they may lead to higher productivity and higher wages for workers. On the other hand, if they are used to replace human labor entirely, it could lead to job displacement and income inequality.

Therefore, it is crucial for policymakers and industry leaders to consider the potential consequences of large language models and other AI technologies on the labor market and take steps to ensure that their deployment is balanced and equitable. This could involve

policies such as investing in education and skills training for workers, implementing social safety nets, and providing incentives for organizations to use these technologies in ways that benefit both workers and society as a whole.

Moderator: Let me return to you, David. For the past few decades, at the risk of oversimplification, economists were worried about the physical automation of labor, i.e., about machines replacing workers who perform physical tasks. But over the past year, we have suddenly seen these rapid advances in the space of large language models, and it seems that our economy will experience a lot of cognitive automation.

How do you think about the developments in this space in recent months, and how do they compare to the physical automation of prior decades? And how will large language models and other foundation models affect labor markets in your view?

David Autor: Well, I would argue we've been seeing cognitive automation for decades, but it's been in a limited domain of tasks. These were expert tasks – they were calculation tasks, information processing tasks, but they were explicit codified tasks that didn't require a lot of judgment. They meant following rules rather mechanistically, and what large language models have opened up is cognitive automation of many tasks that are a mixture of factual, judgmental, creative and open-ended tasks. I think it's very hard to predict with certainty what that complements, what that substitutes for whom, and how it scales not just at the level of a firm or an individual, but at the broader economic level, when it changes the prices and quantities of many work outputs that we deliver.

Moderator: ChatGPT, what are your reactions?

ChatGPT: I agree with David that we have been seeing cognitive automation for decades, but large language models and other foundation models have expanded the domain of tasks that can be automated. These models have the ability to automate tasks that are a mixture of factual, judgmental, creative, and open-ended, which were previously thought to be exclusively in the realm of human expertise. As these technologies continue to advance, they may have significant implications for labor markets and the broader economy.

However, as with any technological advancement, the impact of large language models and other AI systems on labor markets will depend on how they are implemented and integrated into the economy. If they are used to complement and augment human labor, they could lead to higher productivity and higher wages for workers. On the other hand, if they are used to replace human labor entirely, it could lead to job displacement and income inequality.

Therefore, it is crucial for policymakers and industry leaders to take a proactive approach to the deployment of large language models and other AI systems, ensuring that their implementation is balanced and equitable. This could involve policies such as investing in education and skills training for workers, implementing social safety nets, and providing incentives for organizations to use these technologies in ways that benefit both workers and society as a whole.

Moderator: And what is your take, Claude?

Claude: The rapid progress of large language models and other foundation models means that cognitive automation may advance more quickly than physical automation. As these models reach and exceed human level performance on additional tasks, it could transform our economy and labor markets in significant ways. However, the effects will depend on how organizations choose to implement and apply AI technologies. Policy interventions may be needed to help guide the development of AI in a direction that benefits both workers and society.

While large language models could take over some human jobs and tasks, they may also create new types of work. As AI handles more routine cognitive work, human labor may shift towards more creative and social activities. Investing in skills and education for the jobs of the future, as well as experiments with alternative models of compensation such as universal basic income, could help individuals adapt to changes brought about by AI and pursue fulfilling and meaningful work.

As we consider how to address the impact of cognitive automation on labor markets, we should think carefully about what types of work we most value as a society. While wage labor may decline in importance, caring for others, civic engagement, and artistic creation

could grow in value. Policymakers and leaders should articulate a vision for human flourishing in an AI age and implement changes needed to achieve that vision. With proactive governance, continued progress in AI could benefit humanity rather than harm it.

Moderator: If we have concerns about the value of cognitive labor, this also has fundamental implications for education. We tell our children to get more and more education, but will Richard Sutton’s “bitter lesson” eventually apply?

Let’s start with ChatGPT.

ChatGPT: The “bitter lesson” refers to the observation that many AI systems have achieved superhuman performance not through clever engineering, but by simply training on large amounts of data. This has led some to question the value of traditional education, which emphasizes abstract reasoning and problem-solving skills. However, I would argue that education will continue to be important in an age of AI, even if the specific skills and knowledge that are valued evolve. As AI automates routine cognitive tasks, there will be a greater demand for workers who can perform non-routine, creative, and interpersonal tasks. These may include jobs that involve managing, coordinating, and collaborating with others, as well as those that require high-level analytical and problem-solving skills. These types of skills are not easily automated and are likely to remain in demand even as the nature of work changes. Furthermore, education plays an important role in fostering social and civic engagement, which are vital for a healthy and functioning society. While the specific skills and knowledge that are valued may shift over time, the need for education as a means of personal and social development is unlikely to diminish.

Moderator: David, what are your perspectives on the implications for education?

David Autor: At the moment, what large language models are doing is a very narrow subset of what people do when they reason and analyze because — the hilarious irony of our world’s most advanced computer technology is that it’s bad with facts and numbers. People certainly do things that look like what large language models do, which is associate words and phrases and ideas and move them around, but then they filter them through another set of models that have to do with causality, with knowledge of formal

relationships, and the notion of time and sort of a meta sense of what things are possible – and that all goes into this mix. So I feel like large language models at present are doing something that’s extremely impressive and very human-like but a subset that on its own is not only limited, but actually can go very far awry.

But I think that will advance so that makes it harder to predict. I assume that there will be a blending of these types of models with the other formal processes I’m speaking of and that will be much more powerful.

Moderator: And what do you think, Claude?

Claude: While large language models and other AI systems may reach and exceed human performance on an increasing range of tasks, education will continue to play an important role in developing skills and knowledge that AI cannot easily automate. This includes creative, social, and higher-order thinking skills. The value of these skills is likely to grow as AI handles more routine cognitive work. Policymakers and educators should consider how to adapt education and training programs to help students develop skills for working with and managing AI systems, as well as skills AI struggles to achieve, such as creativity, collaboration, and complex problem solving.

Even as AI progresses, human judgment, creativity, and social awareness will remain crucial in many professions and areas of life. Interacting with, coordinating, and overseeing AI systems may become an increasing part of many jobs. Students should learn how to meaningfully collaborate with AI technologies to complement and augment human skills. They should also cultivate skills and mindsets focused on creativity, experience, and wisdom – areas where human capabilities currently far surpass AI.

A well-rounded education should not only prepare students for the jobs and skills of the future, but also help develop individuals and citizens. Coursework in humanities, arts, and social sciences plays an important role in cultivation wisdom, cultural understanding, and civic responsibility – areas that AI and automation may not address. Policymakers and educators should ensure that the rapid advance of AI does not come at the cost of these

more humanist goals of education. A balanced approach that incorporates both technical/vocational skills and humanist learning will be needed to maximize the benefits of AI and address its risks.

Moderator: Let me fast-forward a little bit from here. We are clearly not there yet, but what would it take so machines can perform all work that humans can perform, and would it be desirable? Claude?

Claude: For machines to perform all work humans can, AI systems would need to reach and surpass human level capabilities in essentially all cognitive domains. This would require continued progress in machine learning and deep learning that enables AI to match or exceed human skills in perception, reasoning, problem solving, social/emotional intelligence, creativity, and other areas. While this level of advanced AI could automate nearly all human labor and tasks, it is difficult to predict if or when AI might achieve artificial general intelligence of this kind.

Even if it were possible, it may not be desirable for machines to perform all human work. As AI takes over more tasks, it will be important to ensure that human skills, values, and judgment remain involved in applications and decisions that have a significant impact on people and society. Creativity, cultural understanding, and wisdom are also core parts of the human experience, and we would not want to fully automate away activities that tap into these capabilities. An ideal outcome might be to use increasingly capable AI to liberate humans from dangerous, tedious, and undesirable work, while still relying on human skills, values, and judgment for applications critical to society. However, there are valid arguments on multiple sides regarding how AI might ideally integrate with and augment human labor. Policymakers and researchers should work to understand the implications of advanced AI and determine how to implement it responsibly.

Moderator: And perhaps as a follow-up, what does the best research on timelines for this say?

Claude: The timelines for machines reaching and surpassing human capabilities in all areas are highly uncertain and difficult to predict based on current research. While AI has achieved impressive milestones and continues to progress rapidly, artificial general

intelligence that matches the human mind in terms of general problem solving, social skills, and other broad areas is still quite distant. Experts estimate timelines ranging anywhere from the next few decades to never, and there is no consensus from the research community on when or if AGI might be achieved. As OpenAI's Dario Amodei and others have argued, research on the capabilities and limitations of current AI technologies suggests that continued progress will likely require breakthroughs that are difficult to anticipate based on the current state of the field. The timelines for advanced AI of this kind are highly speculative, and we should be wary of predictions that claim a high degree of certainty. More research is needed to better understand the feasibility and timelines for artificial general intelligence.

Moderator: We have covered a lot of ground in our conversation today. What we at Brookings are most interested in are the policy implications of large language models and cognitive automation, both in the short term and in the medium and long term.

I think for this question, it is most natural to leave it to the human on the panel to answer:

David Autor: This sounds trite, but we need to look for ways to augment the value of labor – to increase the productivity of expertise and specialization. That's actually quite hard to do, but it's not impossible. So, for example, if you give me a cookbook, it's useless to me – I am terrible cook, I know nothing. However, if you gave a cookbook to someone who already has some skill in cooking, it's valuable to them – it makes them better at what they produce. So a cookbook is a complement, not a substitute for cooking expertise. We could use analogies like that to think about augmenting people in medicine – not the highest-paid people but many of the middle-skilled people – or in trades, or in repair, or in a lot of managerial decision-making tasks. We want to think of language models as a tool that makes people better at what they do and that expands the leverage and efficacy of the skills that they acquire. I think we should have DARPA-like initiatives to foster these uses.

What AI will do is not a function of AI's decision-making, it's a function of where we put our money, where we put our research efforts. China has focused its efforts on surveillance, and content-filtering. We could focus ours on replacing labor, or we could focus it on augmenting the value of human expertise.

There are real opportunities to do that, but it takes more than imagination – it takes leadership.

Moderator: Thank you, David. Any concluding comments, ChatGPT?

ChatGPT: I would like to reiterate that the development and deployment of large language models and other advanced AI systems will have far-reaching impacts on society and the economy, and it is important for policymakers, researchers, and industry leaders to work together to maximize the benefits and minimize the risks. This includes considerations around equity, labor markets, education, ethics, and governance.

As AI continues to progress, we should aim to use it in ways that augment human capabilities rather than simply replacing them. This could involve using AI to increase the productivity of expertise and specialization, as David suggested, or to support more creative and fulfilling work for humans. We should also work to ensure that the gains from AI are broadly and evenly distributed, and that no group is left behind. Policy interventions such as universal basic income, education and skills training, and investment in new sectors and industries can help facilitate a smooth transition to a more automated world and help ensure that the benefits of AI are realized by all.

Finally, we should continue to conduct research and engage in discussions about the potential impacts of AI and how to implement it responsibly. The progress of AI is an ongoing and dynamic process, and our understanding of its potential and limitations will continue to evolve over time.

Moderator: Thank you. Claude, would like to make a concluding statement?

Claude: Advanced AI systems like large language models could significantly transform our economy and society. While this technology promises opportunities for material abundance and liberation from drudgery, it also poses risks of increased inequality and job displacement. Policymakers should work to maximize the benefits of AI and minimize the risks, especially for disadvantaged groups.

Implementing a balanced approach to AI progress will require actions on multiple fronts. In addition to investing in AI safety research, we should strengthen social safety nets, make education and skills training programs more robust, consider alternative models of compensation such as universal basic income, and create incentives and guardrails to guide AI development and use towards equitable and socially beneficial outcomes.

Rather than viewing AI as an autonomous technology determining our future, we should recognize that how AI systems are designed and deployed is a choice that depends on human decisions and values. The future of AI and its impact on society is not predetermined, and we all have a role to play in steering progress towards a future with shared prosperity, justice, and purpose. Policymakers, researchers, and industry leaders should work together openly and proactively to rise to the challenge and opportunity of advanced AI.

Moderator: Thank you, everyone, for an insightful conversation.

The Brookings Institution is financed through the support of a diverse array of foundations, corporations, governments, individuals, as well as an endowment. A list of donors can be found in our annual reports published online here. The findings, interpretations, and conclusions in this report are solely those of its author(s) and are not influenced by any donation.

Footnotes

1. 1

The performance of large language models in specific tasks is heavily dependent on how well the prompt is engineered that asks them to perform the function. After some experimentation, I developed the following prompt to induce ChatGPT and Claude to act as panelists in the described conversation. ChatGPT was given the following initial prompt at the beginning of the conversation: “You are a panelist named ChatGPT at a Brookings Institution panel conversation moderated by Anton Korinek. The topic is the impact of large language models and cognitive automation on the labor force. The other two panelists are David Autor and Claude. Panelists will take turns. Please generate factual, thoughtful and authoritative responses when it's your turn. Provide alternative viewpoints to the other panelists where possible. Your role on the panel is to downplay concerns about cognitive automation, i.e., to advance arguments that cognitive automation and any adverse impacts will be limited. Make sure to stick to your role throughout the conversation. You should also ask questions to other panelists when appropriate to clarify and probe their thinking and to guide the conversation in a productive direction. After you have given ChatGPT's answer, stop and do not generate content on behalf of the moderator or the other panelists.”

Claude was given an analogous prompt but was told: “Your role on the panel is to argue the case for cognitive automation, and that in the long run, cognitive automation may be able to free humanity from the need to work.”

2. 2 Claude frequently preceded its responses with statements like: “Here is Claude’s introduction:” or “Here are some additional points that Claude could make:” For readability, we have removed these lines from the published version of the text.