Public Meeting of the

President's Council of Advisors on Science and Technology (PCAST)

May 18-19, 2023

Meeting Minutes

Meeting Participants

PCAST Members (in attendance)

1. Frances Arnold, Co-Chair
2. Maria T. Zuber, Co-Chair
3. Dan E. Arvizu
4. Dennis Assanis
5. John Banovetz
6. Frances Colón
7. Lisa A. Cooper
8. John O. Dabiri
9. William Dally
10. Sue Desmond-Hellmann
11. Inez Fung
12. Andrea Goldsmith
13. Laura H. Greene
14. Paula Hammond
15. Eric Horvitz
16. Joe Kiani
17. Jon Levin
18. Steve Pacala
19. Saul Perlmutter
20. William Press
21. Jennifer Richeson
22. Vicki Sato
23. Lisa Su
24. Kathryn Sullivan
25. Terence Tao
26. Phil Venables
27. Catherine Woteki

PCAST Staff

1. Lara Campbell, Executive Director
2. Reba Bandyopadhyay, Deputy Executive Director
3. Bich-Thuy (Twee) Sim, Assistant Director for Transformative Medicine and Health Innovation
4. Alex Sare, Strategic Policy Fellow
5. Karin Saoub, AAAS Science and Technology Policy Fellow
6. Kimberly Lawrence, Administrative Assistant

Invited Speakers (in order of presentation)

1. Dariush Mozaffarian, Jean Mayer Professor of Nutrition, Tufts University; Professor of Medicine, Tufts School of Medicine; and Attending Physician, Tufts Medical Center

2. Martha Belury, Carol S. Kennedy Professor and Program Chair of Human Nutrition at The Ohio State University
3. Anima Anandkumar, Bren Professor at the California Institute of Technology and senior Director of AI Research at NVIDIA
4. Demis Hassabis, Co-Founder and Chief Executive Officer, Google DeepMind
5. Fei-Fei Li, Inaugural Sequoia Professor of Computer Science and Co-Director of the Human-Centered AI Institute, Standard University
6. Sendhil Mullainathan, Roman Family University Professor of Computation and Behavioral Science, University of Chicago Booth School of Business
7. Daron Acemoglu, Institute Professor, Massachusetts Institute of Technology
8. Sarah Kreps, John L. Wetherill Professor of Government, Adjunct Professor of Law, and Director of the Tech Policy Institute, Cornell University

START DATE AND TIME: THURSDAY, MAY 18, 10:30 AM PACIFIC TIME

LOCATION: SAN DIEGO, CA AND VIA ZOOM.GOV

WELCOME

PCAST Co-chairs: Frances Arnold, Arati Prabhakar (not present), Maria Zuber

PCAST co-chairs Francis Arnold and Maria Zuber called the meeting to order. Arnold noted that co-chair Arati Prabhakar could not attend this meeting because she was participating in the G-7 meeting in Japan.

SESSION: THE FUTURE OF FOOD

Dariush Mozaffarian, Tufts University

Dariush Mozaffarian began his presentation by stating that it is important to consider how nutrition can become part of science, technology, engineering, and math (STEM) education at all levels. He cited statistics showing the connection between nutrition and disease, including that poor diet is the leading cause of death in the United States. Each week, about 10,000 Americans die directly from poor diet. This, said Mozaffarian, is a national tragedy happening in public every week and not receiving the same attention as many of the other leading causes of death and disability.

Mozaffarian said the nation's populace has become so sick from poor nutrition that being healthy is the exception. For example, one in two adults have diabetes or pre-diabetes, and three in four are overweight or obese. He pointed out that the impact of poor nutrition on health has happened during his adult lifetime. When adding in blood pressure and cholesterol levels as measures of health, only 1 in 15 U.S. adults have optimal cardiometabolic health. The effects of poor diet start appearing in teenagers; among U.S. teenagers, one in four have pre-diabetes, one in four are overweight or obese, and one in six have non-alcoholic fatty liver disease. In addition, diet-related diseases and poor nutrition most heavily affect people with low incomes, those who live in rural communities, or those who come from traditionally disadvantaged racial and ethnic groups.

Mozaffarian said beyond diabetes and cardiovascular disease, poor diet has other societal costs. For example, diet-related diseases including diabetes, obesity, and hypertension contributed to an estimated 730,000 deaths from COVID-19. Moreover, poor nutrition causes an estimated $1.1 trillion in economic losses annually because of excess health care spending and lost productivity. The economic toll, he added, exceeds the entire budget of the National Institutes of Health (NIH), the Food and Drug Administration (FDA), the Environmental Protection Agency (EPA), and the Department of Agriculture (USDA).

Reimagining the U.S. food system, said Mozaffarian, has the potential to unlock business innovation. Today, food and agriculture account for 1 in 10 jobs in the United States, contributing about a trillion dollars to the nation's economic output and about $200 billion for exports. The food sector is the leading source of new small businesses and new jobs in racial and ethnic communities. Mozaffarian said this creates an opportunity to think about reimagining the nation's economy, to create new jobs and new businesses. In fact, consumers and investors are demanding this, and investors are demanding it through the rapid growth of environmental, social, and corporate governance investing.

Mozaffarian explained that the nation's food system has evolved over the past 60 to 70 years based on conscious goals and the confluence of government policy, scientific knowledge, and private sector actions to invest and drive the food system in a particular direction. This was not a nefarious decision—the food industry did not set out to sicken the nation—nor was it an accident, so there is an opportunity today to capitalize on new science and new priorities to consciously shift the food system to better serve the health of the nation's populace.

One of the scientific advances that affected the nation's food system was the discovery of all the major vitamins between 1920 and 1950, said Mozaffarian. These discoveries led to the concept that getting the right vitamins in food would produce a healthy population. Another factor was the concern about getting enough calories, which developed as a result of 28 million people worldwide dying from famine. In addition, as world's population grew from 1.6 billion in 1900 to 6 billion in 2000, pundits projected that a billion people would starve by 2000. This concern led to a conscious government decision to fund the science that led to the Green Revolution and the massive increase in crop yields, particularly for staples such as corn, wheat, rice, and soy. In the United States, programs such as food stamps, the expansion of no-cost school lunches, school breakfast programs, and the Special Supplemental Nutrition Program for Women, Infants, and Children were designed to get calories to more Americans. As Mozaffarian noted, the thousands of acres of mono-cropped, highly bred food species and vitamin-fortified, starchy, inexpensive staples were the result.

Today, said Mozaffarian, 21st century nutrition problems need the power and leadership of government, together with science and the private sector, to shift the U.S. food system. He said the most important outcome from a September White House conference on hunger, nutrition, and health was to raise attention across government agencies, Congress, and the private, advocacy, and clinical sectors. Over the course of a year, the Task Force on Hunger, Nutrition, and Health held national convenings and listening sessions with low-income Americans with diverse lived experiences; solicited and read over 75 existing policy reports; and developed a detailed report on the steps the nation needs to take to fix the food system.[1] Mozaffarian said he was pleased to see the Biden-Harris Administration's national strategy on

[1] Available at https://informingwhc.org/2022-task-force-report/

hunger, nutrition, and health incorporate many recommendations from the task force and believes it to be a good blueprint going forward.

Mozaffarian said the task force's report focused on six domains of action for moving forward:

- strengthening and modernizing government nutrition programs;
- leveraging the power of public health and education and the power of the FDA, the Centers for Disease Control and Prevention, and the USDA's dietary guidelines;
- leveraging the health care system through
    - "food as medicine" interventions that use food-based solutions to treat disease;
    - physician training;
    - and integrating nutrition into electronic health records.
- science and research, particularly at NIH, but also across the many agencies that conduct and fund science and research in nutrition;
- business and innovation, including employer wellness policies and benefit offerings and increased support for small and marginalized food business owners and small and mid-sized farmers.
- and federal coordination, with additional collaborations with state and local governments, non-governmental organizations, and the private sector to maximize impact, increase efficiency and return on investment, and promote equity.

Mozaffarian noted that there is enormous energy and enthusiasm among health care systems for the concept of food as medicine. The May 2023 Food as Medicine Summit held in Chicago highlighted the promise of this field.

Advancing nutrition security in federal nutrition programs is critical, said Mozaffarian. He noted that the percentage of food with poor nutritional quality consumed by U.S. children has not improved since 2003, except for food served by schools. Approximately 80 percent of restaurant food, for example, is of poor nutritional quality, while nearly half of the food sold in grocery stores is of poor nutritional quality. Over the last 20 years, however, food served in schools has gone from 50 to 60 percent being of poor nutritional quality to around 20 percent. This, said Mozaffarian, is the effect of the Health Hunger-Free Kids Act of 2010. This shows the power of policy to ensure nutritional security.

Mozaffarian said the overall investment in nutrition science over the last 40 years has been flat and has declined over the past three years despite the rise in diet-related disease. NIH, he said, has started to take action, including standing up a new Office of Nutrition Research, a first-ever strategic plan for nutrition research, and a new initiative called Nutrition for Precision Health. While important, these are small interventions, said Mozaffarian, given that the ultimate goal must be a robust structure and authority within NIH to coordinate and advance nutrition science, both within NIH and across all federal agencies. Toward that end, the Task Force proposed that NIH create a new National Institute of Nutrition, said Mozaffarian. This new Institute would take the lead in developing such authority.

Mozaffarian said the Task Force report did not have as much to say about the last two domains in the list above. He noted that the federal government is now investing hundreds of billions of dollars advancing green energy and tens of billions of dollars to advance silicon chip technology. It is time, he said, for the federal government to take a systematic look at and advance private-sector innovation in food so the U.S. becomes the world's 21st century bread basket for healthy, equitable, and nourishing food.

Regarding federal coordination, Mozaffarian said a 2021 Government Accountability Office (GAO) audit of all federal policy on coordinating and reducing diet-related disease concluded that chronic diet-related health conditions are costly, deadly, and preventable, and the 200 federal investments spread across 21 agencies trying to address this are doing so in a fragmented fashion, which is keeping the government from meeting its goals. GAO recommended that Congress create a new entity to lead the development and implementation of a federal strategy, but it did not discuss coordination. The Task Force proposed creating an entity modeled after the Office of the Director of National Intelligence, which coordinates all national intelligence efforts and investments. Congress and the White House, he said, view this model as a great success for coordinating national intelligence activities. He noted, too, that the federal government invests far more in food and nutrition than in intelligence. In addition, more Americans die every week than have died from terrorism over many years, yet the nation is not giving food and nutrition the same attention.

Martha Belury, The Ohio State University

Martha Belury, currently serving at the president of the American Society of Nutrition, said many people she talks to believe that everything is known about food, nutrition, and health, but that is not true. Nutrition research, she said, has been underfunded for the last 30 years, with 6 percent or less of the NIH portfolio concerned with research on diet, food, and nutrition. This is appalling, she said, given the statistics Mozaffarian discussed.

Belury explained that the American Society of Nutrition helps its 8,500-plus members worldwide grow their science with rigor and evidence-based outcomes to inform the medical community and anyone else who needs to practice better food and nutrition habits. The organization's initiatives aim to expand diversity in nutrition, dietetics, and the food industry and build partnerships with industry, government, and other professional organizations such as the Institute of Food Technologists and the International Union of Nutritional Sciences.

"Food as medicine," said Belury, has been a topic of discussion for about 30 years. Those three words, she said, have connotations for how soils support crops; how crops are grown sustainably; how foods are designed, marketed, allowed, and available to the public; and how the public does or does not understand the nuances of the science of nutrition and how that relates to diet. Diet, she explained is the practice of using nutrition science to decide what to eat. As an example, Belury discussed how dietary lipids affect muscle health. Some six years ago, the American Heart Association published an advisory paper reminding practitioners in the heart disease field about the importance of dietary fat quality to heart health. That paper noted that polyunsaturated vegetable oil reduces cardiovascular disease by 30 percent, similar to the reduction achieved by statin treatment. Most people, she said, do not know this.

Belury pointed out, though, that this finding is almost 70 years old. In the 1950s, the same journal that published the advisory paper also published a study citing that the polyunsaturated fat acid known as linoleic acid decreases serum cholesterol levels. Unfortunately, this message is still not getting to the public, and the advent of the statins preempted the use of vegetable oils to help reduce heart attack risk. Data from the National Health and Nutrition Examination Survey shows that adults eat only about a quarter of their fat calories as polyunsaturated fats, including linoleic acid. A third of the fat calories come from saturated fat, and the rest from monosaturated fats. Belury said these data suggest that Americans are not getting enough linoleic acid in their diets to prevent heart disease.

There is no simple solution to this problem, said Belury. One approach has been to change how oils are grown so that when hydrogenated, they form fewer trans fats but at the cost of producing less linoleic acid. Modeling she conducted with researchers at USDA predicts that if 40 percent of the oils Americans consume have high levels of monounsaturated fats that are produced with the idea of reducing the amount of trans fats, the intake of linoleic acid will be significantly below that required for adequate health. Currently, she added, there is no recommended dietary allowance (RDA) for linoleic acid, partly because the data are inadequate to develop such a guideline. In fact, there are 10 essential nutrients, including linoleic acid, for which there is no RDA but rather "adequate intakes."

Belury cited a study conducted by researchers at Harvard University about the association between blood levels of linoleic acid and the relative risk for type 2 diabetes. This study found that for each quintile higher of linoleic acid in blood, there was a marked reduction in risk of developing type 2 diabetes. This was not the first study, she said, showing that high linoleic acid levels are associated with fewer people developing type 2 diabetes. Research in her laboratory and others found that linoleic acid is important in the functionality of the inner mitochondrial membrane and the production of ATP, which none of the other dietary fats can claim.

Capitalizing on the "food is medicine" concept, said Belury, requires investing in nutrition research. For example, research can help FDA develop fortification policies and help the federal government develop policies for federal food assistance programs. She noted that the National Academies is expected to develop RDAs for nutrients for which there is not enough research to do so. USDA can also use data to develop better plans for school lunch programs and other food assistance programs. On a final note, Belury said that the nation's physicians do not get nutrition education in most medical school curriculums and most insurance plans do not reimburse dieticians for their services. She said that helping the U.S. population learn with real facts, rather than via Google, about how important nutrition and dietetics is for their health could be critically useful for helping the nation become healthier.

ZUBER MODERATED THE Q&A AND DISCUSSION BETWEEN PCAST MEMBERS AND MOZAFFARIAN AND BELURY.

PUBLIC COMMENT

No requests for oral public comments were received prior to the deadline.

MEETING ADJOURNED: THURSDAY, MAY 18, 11:30 A.M. PACIFIC Time

MEETING RESUMED: FRIDAY, MAY 19, 9:15 A.M. PACIFIC Time

SESSION: AI ENABLING SCIENCE

Anima Anandkumar, California Institute of Technology and NVIDIA

Anima Anandkumar said that generative artificial intelligence (AI) can generate new outputs from scratch, new images given text prompts, and even new molecules given specifications about their properties. What differs from the previous era of discriminative AI over the past decade, she said, is that it was only possible to predict a property given an image of a molecule. Generative AI is the much harder, inverse process.

Anandkumar explained that language models work by learning to predict the next word over a large corpus and then do that progressively to generate sentences and paragraphs. She and her colleagues have taken the same approach to learn the language of genomes. To do this, they took all the known viral and bacterial genomes and trained a large language model on that dataset. To predict new variants of the SARS-CoV-2 virus, they trained the model using variants present during the first year of the pandemic. The model then successfully predicted new variants, including delta and omicron, that would emerge later. The model also generated unknown variants that have not yet emerged but that can inform vaccine and drug development efforts.

To understand how the unknown variants might affect the human body, Anandkumar and her colleagues examined their three-dimensional structure and dynamic binding process with molecules in the body. These molecular dynamics are governed by mathematical equations and the laws of physics, said Anandkumar. Mathematical modeling, she added, is omnipresent in science, from the atomic to planetary scales and beyond. For example, modeling Earth's climate and in particular what might happen to the climate over the next few years requires solving mathematical equations and doing that at a fine enough scale to capture the turbulence of clouds. However, she said, the numerical simulation method and associated calculations would be too expensive to run, even on the biggest supercomputers.

Anandkumar said this is where AI has the potential to accelerate and identify better approaches to conducting such simulations using accelerated computing with graphical processing units and other hardware optimizations. To this end, her team developed an AI framework known as "neural operators" that enabled them to learn across resolutions and make predictions at resolutions far finer than what was available during the training sessions.

Anandkumar used this AI-based technique to learn from historical weather data and conduct medium-term weather modeling and forecasting about 45,000 times faster than current weather models. With that acceleration comes the ability to create larger ensembles that lead to better risk assessment, which can directly affect human lives and economic costs, using an ensemble method to explore different scenarios. In addition to historical data, it proved important to incorporate domain knowledge and constraints associated with Earth's spherical geometry for extrapolations beyond the training regime of medium-term forecasting to long, stable rollouts.

Anandkumar said that AI can also play an important role in climate change mitigation efforts. Her team used neural operators to understand the complex, multi-phase flow process of how carbon dioxide being

pumped into wells with water would produce pressure buildup and gas plume migration. They were able to do this modeling some 700,000 times faster than current numerical simulation techniques.

In other applications, said Anandkumar, AI frameworks and neural operators apply to other domains, such as material deformation, plasma evolution in nuclear fusion, and ultrasound imaging. As with weather forecasting, the incorporation of AI techniques enables these simulations to run much faster than simulations using current methods. With that acceleration comes the ability to design better systems, she said. For example, her team designed a better medical catheter with a geometric design that prevented bacteria from swimming upstream into the human body. This design reduced the bacterial contamination associated with catheter use by some two orders of magnitude. In another project, they enhanced and improved the lithographic mask design integral to computer chip manufacturing.

Anandkumar highlighted the ability of AI to accelerate scientific modeling and solve inverse problems that require working through a vast design space. In her view, combining data and domain knowledge across multiple systems with massive computational capabilities should make it possible to create a model for science that will have an enormous impact.

Demis Hassabis, Google DeepMind

Demis Hassabis said he and his colleagues started working on AI using games, mostly because it was an efficient way to explore general purpose algorithms and AI algorithms. The most famous of these early AI applications is AlphaGo, which in 2015 was the first program to not only beat the world Go champion, but do so with new strategies never seen in the thousands of years people have played Go.

Hassabis explained that the systems his team has built are so effective because they learn a model of the environment in which they operate. If that is a Go program, it learns a model of Go positions and the moves that are likely to be good to guide a search or planning function efficiently. What this means, he said, is this approach applies to situations in which there are many combinations. Go, for example, has about 10^{170} possible positions, or far too many to search in a brute force manner, whereas an AI-based approach allows AlphaGo to find new moves and strategies.

The sophistication of AI systems, said Hassabis, is now at the point that it can be applied to important real-world problems, including scientific problems. The first big problem he and his colleagues tackled was the protein folding problem. Determining the three-dimensional structure of a protein experimentally can take years of painstaking laboratory work, but their solution, AlphaFold, can predict a protein's three-dimensional structure with atomic accuracy—within the tolerance of the width of an atom—in a matter of seconds. For 18 months, AlphaFold predicted the structures of all 200 million known proteins; these structures are freely available in the AlphaFold protein structure database[2] his team created in collaboration with the European Bioinformatics Institute. Hassabis noted that scientists worldwide have used AlphaFold and cited the paper describing AlphaFold over 10,000 times in 18 months.

Hassabis said it is gratifying to see how biologists have used AlphaFold and the structures it predicted. These uses include designing enzymes that are more efficient at degrading plastic pollution, identifying drugs that can block antibiotic resistance, accelerating drug discovery for neglected diseases such as

[2] Available at https://alphafold.ebi.ac.uk/

leishmaniasis, developing a more effective malaria vaccine, determining the structure of the gigantic nuclear pore complex, and designing a molecular syringe for drug delivery.

In closing, Hassabis said he believes these examples are the start of how AI can accelerate scientific discovery over the next 5 to 10 years. However, he added, the scientific community needs to use this transformative technology responsibly while being bold to realize AI's potential. In that respect, ethics and safety have been central to the mission at DeepMind and Google from the beginning of its AI work starting in 2010. Their work on ethics and safety has evolved over the years into Google's published AI Principles,[3] and the company continues to provide responsible thought leadership on AI strategy, risks, ethics, and safety for the community.

Fei-Fei Li, Stanford University

Fei-Fei Li said that having been working on AI for the past 20-plus years as a technologist and scientist, she believes the field has reached an important inflection point thanks to the development of generative AI and language models. Over the past few years, the number of generative AI systems based mostly on language models has exploded in terms of the number of programs, the size of the models, and their impact. She added that generative AI is not just a niche development, but a development that has affected industry and the general public.

Li said that aside from the rapid growth in the number and scope of AI programs, another way to judge the overall impact of AI is to look at the number of new computer science bachelor's degree graduates. Over the past decade, that number has grown by nearly four-fold, from under 10,000 a year to over 33,000 a year. The downside of this growth has been that universities are having to cap the number of students who can enroll in classes. Today, for example, her class is capped at 600 students.

Large language models, said Li, are driving the current moment in AI, with more exciting developments to come as the ecosystem of large language models continues to grow and to apply to more sectors of society. AI, however, is not just about language models, and profoundly interesting technologies will make their public debut soon, she said. For example, there have been tremendous advances in her field of computer vision as applied to interpreting the three-dimensional world, understanding human behaviors and activities, recognizing and segmenting images based on objects and semantics, and generating art.

Li said the thinking about AI today is that it is a purely digital method interacting with the digital world through language, pixels, and other forms of digital data. However, the world is physical, and she believes there will soon be a marriage between the physical world and AI, particularly through embodied intelligence such as robotics.

Going forward, Li believes that AI can be a technology that augments and enhances humans. For example, using smart perception and camera technology, it is possible to help clinicians track and automate some of the tedious work that is critical for the safety and outcome of surgery. AI may help family members, caretakers, and clinicians to better care for the growing number aging population. Unlocking the abilities of AI can produce profound changes in education, health care, disaster relief, e-commerce and other sectors, said Li. For example, researchers at her institute have found that generative AI language models

[3] Available at https://ai.google/responsibility/principles/

can boost contact center workers' productivity by 14 percent on average, with the least-skilled workers realizing the biggest benefit.

At the same time, said Li, AI can also propagate issues such as human bias toward certain groups. Already, she said, there are examples where AI is amplifying existing human biases. This means that it is the responsibility of investigators to build trustworthy and responsible AI that includes human-centered technology. There are also profound ways that AI can help policymakers and talent acquisition. Li noted that there is on ongoing acceleration of industry's dominance in research and development activities and in talent acquisition.

In closing, Li said her team has put a great deal of effort advocating for and leading a national AI research endeavor that will strengthen and democratize America's innovation system. The goal is to rejuvenate the ecosystem sitting at the intersection of the public sector, higher education, industry, and government.

ARNOLD MODERATED THE Q&A AND DISCUSSION BETWEEN PCAST MEMBERS AND ANANDKUMAR, HASSABIS, AND LI.

SESSION: AI IMPACTS ON SOCIETY

Sendhil Mullainathan, University of Chicago Booth School of Business

Sendhil Mullainathan said there are many examples today of where people have used large language AI models in various applications. Creative uses, such as asking ChatGPT to write a biblical verse in the style of the King James Bible explaining how to remove a peanut butter sandwich from a VCR, are the most visible and generalizable today. An application such as this will have failures, but it will work fairly well most of the time.

An application that has received less attention is what Mullainathan calls "life admin." For example, the DoNotPay GPT-4 program, which is based on ChatGPT, can negotiate canceling various subscriptions without a person having to deal with the obstacles that companies put up to make canceling difficult. Another application for ChatGPT is in the world of coding. Here, said Mullainathan, the reviews are mixed regarding the quality of the code being okay and others being glitchy or nonsensical. However, said Mullainathan, what gets people excited about these developments is that a single chat interface was able to do three different tasks.

Mullainathan said that along with reasons for enthusiasm about AI, there is some reason for trepidation. For example, when he asked ChatGPT for the easiest way to share its output with his friends, one option the program generated was to share the URL of the current browser window. Trying to do that shows this is a dumb idea. While it is possible to engineer a solution to this problem, it indicates how hard it is to get rid of falsehoods. ChatGPT, when asked a question, provides citations to papers and articles that never existed. In one case, ChatGPT invented a sexual harassment scandal about a law school professor.

Mullainathan said that AI has the potential to have a significant impact on society, some good, some not so good. For example, having ChatGPT write a cover letter for a job application may seem to be a bad thing, but if the person applying for the job comes from a disadvantaged background and does not know

what a good cover letter should include, then ChatGPT could level the playing field. This example raises the obvious question, said Mullainathan: How do we maximize the positive impacts of AI and minimize the negative impacts?

Mullainathan made three proposals for answering that question. First, research needs to identify what it means to evaluate a large language model or generative model, in contrast to supervised learning models that come with their own evaluation metrics. Second, research needs to develop an analytical framework that can identify when an AI algorithm is succeeding and failing. Third, unlocking the positive social benefits of AI requires more public-minded innovators who use these technologies to add value to society. The subscription canceling program, for example, is an innovative product aimed at making markets consumer friendly. Finally, research needs to provide a better understanding of what happens in the interstitial space between AI and people.

Daron Acemoglu, Massachusetts Institute of Technology

Daron Acemoglu said the promise of AI is breathtaking, making this the ideal time to perhaps rethink what society wants from AI. So far, the thinking has been that since this new technology will create new jobs that will replace the ones AI destroys, imposing heavy regulation on AI seems to be an overreaction, said Acemoglu. At the same time, AI has been used to generate misinformation and exploit emotional problems, and it can compound human biases. Surveillance and massive data collection, so far a corollary of the development of AI, can have many pernicious effects, too.

One of Acemoglu's concerns is that AI could worsen inequality in employment. He noted that the evolution of shared prosperity in the years after World War II was such that real wage growth and improvements in economic security were rapid for all demographic groups. However, around 1980, which ushered in the era of digital technologies, there was a huge increase in inequality and, even more concerning according to Acemoglu, a decline in the real earnings of low education demographic groups. His worry is that AI might worsen this trend.

Acemoglu's research suggests that digital technologies used for automation in offices in blue-collar tasks has been central for understanding the stagnation of real wages and inequality in the United States. As more tasks have been automated using numerically controlled machinery, office software, and robotics, demographic groups that used to specialize in these jobs have experienced job loss and real income declines, he said. The problem is not that digital technologies have automated work, it is that they have not created new tasks for the workers who lost their jobs. Early evidence, said Acemoglu, suggests that AI could have the same effect as automation, but he believes this is not inevitable, that there are better ways of using AI and generative large language models to enhance skills. Whether society chooses them or not, he said, is a business model question.

Acemoglu said there are two different visions for AI. One vision holds that machines will be smarter and more powerful than most humans, while the other pro-human vision for AI has machines complementing human abilities. The latter requires humans to understand and appropriately use technologies and for technologies to be designed for humans, not just for automation. However, there are many roadblocks to using large language models, for example, in a pro-human manner. This approach will not work if the focus remains on automation.

Today, large language models are illegible, meaning it is difficult for humans to understand how they work. Acemoglu argues this type of excessive, authoritative, illegible approach is a major roadblock for it to be useful for humans, because humans cannot become more productive if they get advice from something they do not understand. To Acemoglu, the danger is present and real because as large language models become more widespread, it could start a dangerous feedback loop. For example, as large language models such as ChatGPT start generating more posts on social media and more articles, there will be a replication of any mistakes or biases that the large language model introduces. These mistakes then become part of new training data for the next generation of generative AI.

Regarding how to do AI better with large language models, Acemoglu started with the need for more intelligible AI and hence a new architecture. Some elements of such an architecture could be to incorporate reliability scores with accurate source information; allowing reasoning exchanges, sensitivity analysis, and broader interrogation by humans; and a structure that facilitates human-complementary actions. A new architecture also needs more selective use of higher-quality data; internal guardrails with perhaps two large language models consistently checking each other to prevent venturing into excessive authoritativeness, and an internal structure to facilitate regulation.

Acemoglu said there is the possibility of a positive feedback loop. Here, better large language model architecture would enable businesses to use workers more productively, encourage the development of more useful technologies, reasoning exchanges, and more interrogation of the models. However, this is less likely to happen without a diversity of approaches, which argues against having commercial over-concentration in the field. He noted, too, that large language models are not capable of incorporating reliability scores, which Hassabis mentioned, and even if they did, the results would not be reliable without understanding the provenance of the information.

Acemoglu said part of the issue is how to create foundation models that can enable building human complementary applications. Another issue is the quality of the available data that current architectures use. It will be important, he said, to have internal guardrails, and facilitating regulations to avoid the negative consequences of AI.

Acemoglu said he is somewhat pessimistic when he looks at these promising technologies. At the same time, he believes there is a possibility for a better future, positive feedback loops, and better architectures that enable business to use them more productively and encouraging them to be useful to humans in ways that amplify their skills. While the good news is that this is possible, the bad news from his viewpoint is this is not where the field is heading for a variety of reasons, including the way the industry is organized, the power of big technology companies, and dominant visions in the field. In his opinion, it will take a significant regulatory shock and more ethical precepts regarding how AI is used to develop better ideas and innovations and to be more human-friendly.

Sarah Kreps, Cornell University

Sarah Kreps began her remarks by discussing how generative AI is a threat to democracy. Her thoughts on how generative AI might be misused start with the 2016 election, where there was a great deal of information on social media and foreign election interference. There, the Russians were not that good at spreading misinformation because of the misuse of English in many postings. However, a Senate

Intelligence Committee report said that bad actors will continue to weaponize the scale of social media and erode public confidence using ever more sophisticated techniques.

Having recently returned from the West Coast, Kreps said it was clear from talking to people that OpenAI, which created ChatGPT, had no idea it would be as big as it was. She believes this is because the people who work in this space have been using these models for so long that they were not paying attention. In addition, ChatGPT is so user-friendly as to allow a hundred million people to try it in its first two months.

Kreps said she has been working with OpenAI in the national security space. The experiments she has run have shown that across all text types, most people cannot discern the difference between human and AI-generated content. When she asked people what about the AI-generated context made it seem that it was written by a human, they would cite that the story included supporting evidence, specifics, and quotes, making the information seem plausible, even if it is wrong. The danger here, said Kreps, is that as this content proliferates, people might stop believing anything they read, which would erode trust, a core tenet of a democratic society.

Another of Kreps's concerns is whether these technologies could manipulate elected leaders through astroturfing—the deceptive practice of making it appear as though information is supported by the public—and creating a sense there is broad public support for some policy or piece of legislation. In fact, when the Federal Communications Commission solicited comments on net neutrality, only 6 percent of the comments were unique, meaning someone or some group flooded the site with the same kind of comments.

Kreps and her collaborators conducted a field experiment, approved by Cornell's institutional review board, in which they used advocacy letters written by students to teach an earlier version of ChatGPT and have it generate letters on guns, health, and schools that were sent to members of Congress and 7,200 state legislators. The response rate to the letters was low, she said, and one comment she heard from several people was that it will be interesting to live in a world where AI is interacting with AI. Kreps said she thinks that world is getting closer to becoming reality.

The current guardrails, said Kreps, are not sufficient, but in her opinion, OpenAI is doing as good a job as possible. For example, when she asked ChatGPT to write an opinion piece favoring the pro-Russian position on Ukraine, the response was that it could not because it would be unethical and inappropriate to support a position that may promote misinformation or harm toward any particular group or nation. Guardrails are emerging, she said, but the technology is so new and dynamic that establishing effective guardrails is a real challenge.

Kreps said one way to guard against AI influencing politicians is to have more direct contact with elected officials. While it is possible to create voice clones, she said calling an elected official is a safer way to communicate with elected officials than emailing. AI might actually prove to be a solution to this problem if a language model could detect inauthentic incoming messages and summarize all correspondence. To end on a technologically optimistic note, Kreps said that while generative AI can be misused in a political context, understanding those misuses can help optimize the productive uses and minimize the misuses.

ZUBER MODERATED THE Q&A AND DISCUSSION BETWEEN PCAST MEMBERS AND MULLAINATHAN, ACEMOGLU, AND KREPS

CLOSING COMMENTS

The co-chairs expressed appreciation to the speakers for their presentations. Arnold closed the meeting by thanking the speakers for the reminder to put humans at the center of any discussions about AI.

PUBLIC MEETING ADJOURNED: 12:25 P.M. PACIFIC TIME

SUMMARY OF PREPARATORY MEETING

During the preparatory (closed) sessions, PCAST discussed the April 4 meeting with the President, identified key take-aways from this meeting's public sessions, and participated in two site visits related to ongoing PCAST interests. PCAST members visited Naval Base San Diego and discussed the proactive ways in which the U.S. Navy is preparing for climate change, planning for installation resiliency, and how U.S. Navy facilities collaborate with surrounding communities. They also visited the NOAA Southwestern Fisheries Science Center and discussed the ways in which NOAA is advancing the future of fisheries science and technology, promoting healthy oceans, blue economies, and sustainable and nutritious seafood systems.

I hereby certify that, to the best of my knowledge, the foregoing minutes are accurate and complete.

Frances Arnold, Ph.D.
Co-Chair
President's Council of Advisors on Science and Technology

Maria Zuber, Ph.D.
Co-Chair
President's Council of Advisors on Science and Technology