

ACCELERATE

Co-Developing Accessible Advanced AI

Abstract

There are 56 artificial intelligence (AI) startups worth over \$1 billion today. That is a testament to the enormous power of deep learning, which has found transformative applications in everything from finance to healthcare. These approaches require huge amounts of data and computational power, however, which means that advances are increasingly driven by a handful of large companies and governments. We are about to enter a “third wave” of AI that will imbue machines with “common sense” and reasoning capabilities, allowing much broader deployment, and increasing the breadth and depth of human-machine interactions. That makes it crucial that these advances are not shaped by narrow interests and that everyone can take part in the development of advanced AI and benefit from its use.

- What will the next generation of AI look like and how should we best prepare for it?
- What priorities should inform the next stage of AI development?
- How will advanced AI be able to address global challenges differently than today’s technology?
- What can we do to avoid “AI nationalism” and ensure broad access to the technology and applications developed on the basis of advanced AI?

Participants

Moderated by:

Amandeep Gill, Director I-DAIR project, India

With:

Pushmeet Kohli, Head, AI for Science, DeepMind, India (*remotely*)

Nanjira Sambuli, Policy Analyst, Advocacy Strategist; Board Member, Digital Impact Alliance, Development Gateway, and The New Humanitarian; Member, GESDA Diplomacy Forum, Kenya

Daren Tang, Director General, World Intellectual Property Organization; Member, GESDA Diplomacy Forum, Singapore

Rüdiger Urbanke, Professor of Communication Theory, EPFL; Member, GESDA Academic Forum, Austria

Wendell Wallach, Senior Advisor, The Hastings Center, United States

Ewan Birney, Deputy Director General, EMBL; Director, EMBL-EBI, UK



Highlights

Artificial intelligence (AI) has many common daily uses including to manage homes and drive cars. Artificial general intelligence (AGI) – the hypothetical ability of a machine to perform tasks equally to humans, with common sense and general reasoning – remains years away. However, advanced AI systems that can process reams of data, learn by themselves and find solutions beyond what most people could discover already signal the potential for huge breakthroughs.

At Google DeepMind, for example, Pushmeet Kohli’s team recently developed AlphaFold, an AI software that, basing its calculations work on extremely huge amounts of public data on proteins, predicts their 3D structure from the amino acid sequence. Considered an AI solution to a half-century “grand challenge” in biology, this rapid system to determine the shape of proteins could pave the way for a host of new uses such as treatments for diseases or enzymes that break down industrial waste. And in 2015, DeepMind created AlphaGo, a programme that combines advanced search tree with deep neural networks and which, for the first time, defeated a professional human playing the Go board game. There are few scientific areas that have not been impacted or completely revamped by the “pervasiveness” of advanced AI such as the breakthrough solutions developed by Kohli’s research teams, said Rüdiger Urbanke, a computer scientist, professor and pioneer in coding theory and communications theory. “Essentially, they are based on neural networks and our ability to process and collect huge amounts of data,” he said. “And so, for the next probably five, ten years, we’ll see still a lot of this exploitation of this paradigm and amazing things will happen.”

It is not just scientists following this fast-developing field with intense scrutiny. Digital technologies also play a key role in the UN’s 2030 Agenda for Sustainable Development, particularly with diversity, equity and inclusion. As advanced AI promises to make the world more accessible, deep scepticism remains with AI algorithms and technology. Experts point out these are culturally constructed, shaped by the experiences of scientists and technologists. New products can radically change how we communicate, interact, and learn, potentially including or excluding more people.

AI algorithms and tools are often developed and used in ways that experts say are discriminatory towards minorities and vulnerable groups. Therefore, more regulation is needed to prevent bias, said Nanjira Sambuli, a policy analyst and researcher who studies how information and communication technology affect culture, entrepreneurship, governance and media. “It comes down to a number of things: whether it’s already obvious that society is not monolithic, who is involved in shaping how

AI advances, or whose worldviews are incorporated, or even which versions of society incorporated into what is training these machines versus who’s been left out. And how do we correct for that?” she asked. To create a fairer world, the worlds of science, diplomacy, business and civic society must question their assumptions towards AI “because without doing so, we’ll end up with these divides”, according to Sambuli. “That is not something we want: conflicting humans and conflicting AI. That’s not a world I personally want to be in.” The rise of AI is moving from data input to context and experience, leading to the “politics of who’s a creator and who’s the subject”, she added. “How do we change the paradigm of inclusion from being this afterthought? You’ve built a table, and you’re just bringing an extra seat, versus building a table altogether that remembers that these seats should have been there in the first place. That should be a key concern for everybody involved in this.”

Another important aspect to consider is accessibility to raw data. For decades, Ewan Birney recalled, he and his scientific colleagues stored data eventually used for DeepMind’s AlphaFold. “That goes to the importance of data and data-sharing in a fair way to create these possibilities,” said Birney, a biochemist who oversees European Molecular Biology Laboratory (EMBL), Europe’s flagship laboratory for the life sciences. “I was very struck by this business of ‘we shouldn’t just bring a chair to the table, we should make the table all together,’” he said as a prominent advocate of open source data-sharing in bioinformatics and science. “It’s really important that we continue to do data-sharing, but do it in a fair and equitable way, and we do it with co-creation across the world at the start.” If the emergence of big data represents a new global “commons” like oceans and space, then GESDA-backed International Digital Health and AI Research Collaborative (I-DAIR) could help protect it. The initiative, overseen by Amandeep Gill at the Graduate Institute Geneva, is building a global platform to foster inclusive, impactful and responsible research uses of that data, and to head off potential future conflicts.

Wendell Wallach, a prominent bioethicist and scholar who focuses on ethical and governance concerns posed by emerging technologies such as AI and neuroscience, agrees. “The technologies often get deployed before we even know what is problematic with them. By the time we do know what is problematic, they are so deeply entrenched, we do very little to reform them,” he said. As Urbanke noted, some of the things that can go wrong are massive unemployment for those without access to education for the new sorts of jobs that will be created; terrorism by people who co-opt future technologies; and the spread of lethal autonomous weapons, bearing high-powered munitions.

That has prompted the need for what experts describe as 21st century approaches to governance, relying more on “soft law” that is not always enforceable but can be applied quickly through standards, practices, codes of conduct and insurance. “Hardly a day goes by without somebody putting in front of me a new scenario of what we need to be concerned about,” said Wallach. “I, like most of you, perceive technology as a source of both promise and productivity. But there’s considerable disquiet – disquiet not only over specific technologies, but disquiet over the overall trajectory of this vast scientific development and technological deployments that we are witnessing. And for me, I look very much at the two sides of it. How can we reap the benefits but without naively overlooking the detrimental societal impacts?” Wallach said he was particularly worried that “learning algorithms are not transparent”, meaning even those who develop them or the experts who work with them “cannot explain how we got from the input to the output”.



To those points, Kohli said the DeepMind teams that came up with AlphaFold and AlphaGo used a multidisciplinary approach that is important to solve major problems. Since AI and machine learning are powerful technologies that can have “many different side effects”, he said, there should be no “complacency” towards their uses; people with diverse sets of interests should be represented when developing projects. “In our DNA, we are sort of thinking about ourselves as an organization which is representative of what we want to achieve in the world and have everyone on the table,” said Kohli. “Our hiring practices, how we go about developing our project roadmaps, have a first-glance element: we think about diversity, think about ethics, think about impact, about any particular project, even before we take the first step in executing it.”

Overall, the potential uses of advanced AI are a “deep concern” among many nations and global institutions looking at the social, technological, cultural, and corporate implications, said Daren Tang, who before taking over one of Geneva’s most prominent international organizations held various

legal positions dealing with trade and industry in Singapore’s government. In today’s climate of rising nationalist populism, authoritarianism, and disinformation, it’s become “very, very tough” to promote multilateralism, he noted, and “any attempts to push big multilateral norm-setting treaties and international treaties and AI, for example, I don’t think it’s going to work. Not for the foreseeable future anyway”. What might work instead? “First, we can come together to talk about soft laws. Standards can be harmonized on these standards,” said Tang. “More and more of us are beginning to broaden our offerings to do these soft approaches.” That requires an international city like Geneva where “GESDA is a classic example of how Geneva is playing to its strengths, bringing the world of science and diplomacy together,” he added. “More of these things need to happen.”

Asked by the audience if “hard laws” for regulating AI might be needed someday beyond merely relying on soft laws, Tang argued for a phased-in approach due to escalating tensions among the major powers. “Norms are only set when there is consensus amongst key countries,” he said, pointing to past global trade pacts pushed by the European Union, Japan and the United States. “Those days are over.” Now, he said, national interests are paramount such as with corporate global taxes. “The challenge for AI and technology is that it’s become geopolitical. It’s become a key part of competition amongst member states. And you don’t have to look too far,” said Tang. “But we can’t give up just because of that. So what? What do we do next? And if we can get hard law, we get soft law. And I think soft law then becomes a way for us to build up towards hard law, because I don’t think the world is going to be like that forever and ever. And so, let’s prepare for the time.”

Wallach, however, proposed a middle way between hard law, which he said tends to become “crystallized” by installing bureaucracies that stifle innovation, and soft law, which has problems of “enforceability”. He suggested two institutional approaches, both of which he said GESDA could be useful in championing. The first would create a global governance network, as proposed by the UN Secretary-General’s Roadmap on Digital Cooperation last year, setting up a “true multi-stakeholder forum working through these problems together”, Wallach said. The second would be a “rapid response” panel of “good-faith brokers” such as GESDA, the Carnegie Council for Ethics in International Affairs and others with international prominence that could “quickly convene the appropriate experts to work together and see if we can speed up our responsiveness to these various kinds of challenges”, he said. “Responses in this on all three levels of engineering, first of all. What can we and can we not engineer? Secondly, ethics. How do we ensure that the deployment of these systems is appropriate. And third, governance. How can we put in place effective means for that?”

Takeaway Messages

The rise of AI is moving from data input to context and experience.

Few scientific areas have not been impacted or completely revamped by the “pervasiveness” of advanced AI, and many things can go wrong if improperly used.

Deep scepticism remains with AI algorithms, data sets and technology, which experts point out is culturally constructed. More regulation may be needed to prevent bias.

Digital technologies and AI condition access to the world for an ever-larger group of people, making inclusiveness, representativeness and cultural biases ever more important. Inclusivity means building a table for everyone to gather round in the first place, not just adding seats.

Open source data and data-sharing in a fair and equitable way is essential to create new possibilities and solve scientific and technological problems.

Science and technology organizations should incorporate people and strategies that reflect diverse sets of interests from the start when developing projects, not as an afterthought.

More reliance on 21st century governance – such as “soft law” that is not always enforceable but can be applied quickly through standards, practices, codes of conduct and insurance – might be useful.

More information

[Session recording on YouTube](#)

[Related interviews: Nanjira Sambuli part 1, Nanjira Sambuli part 2, Daren Tang & Rüdiger Urbanke](#)

[Tweets related to the session](#)

[Related content in the 2021 Science Breakthrough Radar®](#)

[Advanced Artificial Intelligence and related break-](#)

[throughs at five, ten and 25 years: Full breakthrough brief, Deeper Machine Learning, Human-centred AI, Next-level AI, Interdisciplinary AI](#)

[Brain-inspired Computing and related breakthroughs at five, ten and 25 years: Full breakthrough brief](#)