Report from the High-Level Hearing
‘A European Union Strategy for Artificial Intelligence’
**Introduction**

Deployed wisely, AI holds the promise of addressing some of the world’s most intractable challenges. But the significance of its positive impact is mirrored by its likely destabilising effects on some aspects of economic and social life. Finding a policy response to what is undoubtedly ‘the next big thing’ is both urgent and challenging. How can Europe create a thriving AI ecosystem that will maintain it in the race with the United States and China for global dominance? How can social acceptance of AI be fostered? What are the risks and the benefits? Can Europe set global AI quality standards?

On 27 March 2018, the European Political Strategy Centre hosted six leading international experts for a High-Level Hearing on ‘A European Union Strategy for Artificial Intelligence’ to intellectually accompany the European Commission’s ongoing public consultation on Artificial Intelligence. During the Hearing, the experts were asked to address a set of predetermined questions with the knowledge that a full transcript of the Hearing would be submitted as a public contribution to the consultation.

The experts included:

- **Kate Crawford**, Distinguished Research Professor at New York University, co-Founder and co-Director of the AI Now Research Institute, and Principal Researcher at Microsoft Research
- **Rebecca Crootof**, Executive Director of the Information Society Project at Yale Law School
- **Virginia Dignum**, Professor of Technology, Policy and Management at Delft University of Technology
- **Pascale Fung**, Professor at the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology
- **Rose Luckin**, Professor of Artificial Intelligence at University College London
- **Luc Steels**, Director of the Artificial Intelligence Lab of the VUB – Free University of Brussels

The Hearing was moderated by **Ann Mettler**, Head of the European Political Strategy Centre, and **Mario Mariniello**, Digital Adviser to the European Political Strategy Centre.

The full replies can be found in the transcript. As a ‘teaser’, the first section provides a key takeaway from each speaker. These have been selected by the European Political Strategy Centre.
Highlights from the Hearing

Kate Crawford
‘New ethical frameworks for Artificial Intelligence need to move beyond models of individual responsibility to develop real accountability structures for industries and governments as they design and deploy AI. Otherwise, I worry that we run the risk of voluntary high-level ethical principles that can sound impressive but end up being used to substitute for consumer safety, fairness and democratic scrutiny.’

Rebecca Crootof
‘AI systems attempt to optimise some behaviours or achieve some goal, and it is up to human beings to determine what that goal is and what the parameters are on the process for achieving it. […] Because of its strong data privacy laws, the EU occupies a very interesting and potentially really productive regulatory position with regards to AI on the global sphere. […] Proactive regulation is the Goldilocks solution: it attempts to address specific problems, while still allowing space for innovation and creation’

Virginia Dignum
‘Regulation will contribute and be a necessary step for building better AI systems. […] And there is where the public sector – and especially Europe at its level of governance – has a very important role to play in order to ensure that those values which we can define as European values are taken seriously and are central to the development of AI systems.’

Pascale Fung
‘The lack of a homogeneous linguistic market is a huge barrier to data. Both North America and China have the advantage of having a single language data market. Does that mean that Europe should have a bilingual system where English is the lingua franca of user data? Is that even possible? Is that culturally or politically desirable? I do not know. But if we do not have it here, then Europe needs to consider what kind of leadership position it wants and go from there. […] Strengthen your investment in R&D […]; create a single and more homogeneous marketplace and foster collaboration with non-European partners.’

Rose Luckin
‘Embrace AI. It can bring enormous benefits, but as with everything with huge benefits, it brings huge responsibilities because we need to ensure that those benefits are felt by the whole of society, not just a subset. […] The benefits will also come through disruption. […] In order to accelerate adoption, we need to make sure that, as a human workforce, we’re ready to use AI effectively. […] It’s a real job for the education sector not just to educate those who build AI but also educate people to understand enough about it.’

Luc Steels
‘The main things we have to worry about are: overestimation - with the consequence that unwarranted fear holds everybody back – and underestimation - with the effect that we don’t seize the opportunities that are already there with AI. We need realistic assessments and stable infrastructure.’
Structure of the Hearing

During the Hearing, the experts were prompted to reply to six main questions, including a number of sub-questions, which had been shared with the speakers ahead of the event. The questions were drafted by the European Political Strategy Centre for the purpose of stimulating the discussion. The questions provide no indication as to the European Commission's views on the subjects discussed.

Here are the questions:

• **Question 1**: Please state your name and affiliation; please flag any potential conflict of interest (for example, if you are providing consulting services to a client potentially affected by actual or potential initiatives by the European Union in the area of Artificial Intelligence, please state so). Please describe your background and your experience in dealing with Artificial Intelligence, particularly from a public policy perspective.

• **Question 2**: In your own words, what is Artificial Intelligence? Please provide the definition that you would consider the most accurate. Please also share your general views on global trends and, in particular, on the current and future prospect of AI development in Europe.

• **Question 3**: Please elaborate on the potential benefits yielded by Artificial Intelligence to society: citizens, business and public sector. You are welcome to describe concrete AI applications to illustrate your point. In addition, please elaborate on barriers of any nature that could hamper the adoption and development of Artificial Intelligence.

• **Question 4**: Please elaborate on the potential harm that could be caused by Artificial Intelligence. Please state what worries you the most when considering the impact that Artificial Intelligence may have on human beings, identify the sources of highest concern for society and describe the ethical dilemmas that you consider most pressing.

• **Question 5**: Based on your research and professional experience, please identify the main areas of action for public policies aimed to foster the development and adoption of Artificial Intelligence and to mitigate the risk of its potential harm. In your answer, please focus on the European Union: provide your assessment of the current landscape and your suggestions for an effective and coherent European Union Strategy for Artificial Intelligence.

• **Question 6**: In a nutshell: your message to the European Commission – what should (or should not) be done about Artificial Intelligence.
Full Transcript of the Hearing

Welcoming remarks

Ann Mettler: Good afternoon and welcome to this hearing convened by the European Political Strategy Centre on a European Union Strategy for Artificial Intelligence. My name is Ann Mettler and I’m the Head of the European Political Strategy Centre, the EPSC, which is the European Commission’s in-house think tank. I’m joined on my left by Mario Mariniello, Digital Adviser at the EPSC.

This hearing is organised in order to contribute to the European Commission’s internal reflection on Artificial Intelligence. There will be a communication on Artificial Intelligence, which will be published in April, and it will also be among the topics that EU leaders will discuss at an informal summit in Sofia, Bulgaria, in May. I’m delighted that we have such an excellent high-level group of experts with us today to discuss the various issues surrounding Artificial Intelligence and machine learning.

Before we start the hearing, let me briefly give you some instructions. The hearing will last about two hours until 6 o’clock. Each speaker will have a certain amount of time to address each question. Speakers were provided with an extended list of questions ahead of this meeting. The hearing will be on the record and a full transcript of the hearing will be published on the EPSC website. Given the format, I would appreciate it if the audience would be in full listening mode, as no interaction with the speakers during the hearing will be allowed. However, there will be an opportunity for an exchange of views after the hearing is over, from 6 o’clock onwards, over an informal coffee served right outside this room.

Two more announcements: One minute before the time limit of an answer expires, we will signal this by showing an orange card. Once you have used up the maximum allotted time, you will be shown the red card. So once that is shown, we ask you to conclude your remarks and if you take a bit longer in one question, we ask you then to reduce your time accordingly in another question. Also, we are rotating the order in which speakers are called on. So the order will be different for each question so that we all don’t always have to call on the same people starting or concluding a question round. We will now start with the hearing.

Question 1: Introductions

Let me pose the first question: Please state your name and affiliation; please flag any potential conflict of interest (for example, if you are providing consulting services to a client potentially affected by actual or potential initiatives by the European Union in the area of Artificial Intelligence, please state so). Please describe your background and your experience in dealing with Artificial Intelligence, particularly from a public policy perspective. And we start with you, Professor Crawford.

Kate Crawford: Good afternoon, Commissioners. Thank you for inviting me here today. My name is Kate Crawford. I’m a distinguished research professor at New York University where I co-founded the AI Now Institute. I’m also a visiting professor at MIT and a principal researcher at Microsoft Research. I would like to be clear that I’m speaking on behalf of myself as a researcher and a professor today and not representing Microsoft’s position.

For many years, my work has been focused on the social implications of big data, machine learning and Artificial Intelligence. I’ve worked in both academia and in industrial research. And I publish widely on issues relating to algorithmic bias, fairness and due process.

From a policy perspective I’ve worked with various governments including the Obama White House, where I co-chaired one of their AI symposiums and I’ve also advised the UN, the US Federal Trade Commission and the City of New York. I am also a member of the Partnership on AI. And full disclosure: back in 2008, I think, I contributed to the 2020 Agenda for the Australian Government, where I’m from. But that seems like a very long time ago now.
Ann Mettler: Thank you so much. Next is Rebecca Crotof.

Rebecca Crotof: Thank you so much for the opportunity to be here today. My name is Rebecca Crotof. I am a legal academic, and my research focuses on the intersection and the interrelationship between law and new technology. I especially consider when and how law can channel the development of new technology to promote socially desirable aims.

I am currently the Executive Director of the Information Society Project, which is an independent research centre at Yale Law School focused on issues at the intersection of law, technology and society. I am also a research scholar and a lecturer in law at Yale Law School, where I teach a course on ‘Technology Law’ and a course ‘Artificial Intelligence and the Law’. I am recently a member of the Center for New American Security’s Task Force on Artificial Intelligence and National Security. I don’t anticipate that to be a conflict of interest.

Ann Mettler: Thank you so much. And next is Professor Dignum.

Virginia Dignum: Good afternoon. Thank you for inviting me for this hearing. My name is Virginia Dignum. I’m a professor on social Artificial Intelligence at the Delft University of Technology in the Netherlands. I’m also a visiting researcher at University of Melbourne in Australia.

My research centres around the interaction between agents and people in the context of organisations and specifically on the verification and evaluation of those interactions. I’m currently the Executive Director of the Center of Delft University on Design For Values. I’m also part of the Executive Board of the IEEE Ethically Aligned Design Initiative and member of the scientific board of AI4people Initiative from European Parliament and Atomium and on the scientific board of Responsible Robotics Foundation. I don’t anticipate any conflict of interest.

Ann Mettler: Thank you so much. Next is Professor Fung, please.

Pascale Fung: Hello everyone. My name is Pascale Fung. I’m a professor in the Department of Electronic & Computer Engineering and the Department of Computer Science & Engineering at the Hong Kong University of Science and Technology.

I have been in the field of speech recognition and then natural language processing for close to 30 years and I am an elected fellow of the IEEE and the ISCA, which is the International Speech Communication Association. My research interests in the last few decades have been centred on human-machine interactions via speech, language and – more recently – affect: audio emotions and in image, including facial expressions. My more recent work has been focusing on building systems that are empathetic, so that the machines can be more empathetic and helpful to humans.

My interests in the policy area stem from my involvement as an expert on the Global Future Council on AI and Robotics of the World Economic Forum and also as a member of the Partnership on AI. Recently, we have established a Centre for AI Research – short-named CAiRE – to conduct multidisciplinary collaborative research on all aspects of AI at my university, and I’m the Director of that.

Ann Mettler: Thank you so much.

Rose Luckin: Hello and thank you very much for inviting me, it’s a pleasure to be here.

I am Professor at UCL Knowledge Lab, and my area of expertise is in Artificial Intelligence and education. So that’s both in terms of developing systems, using Artificial Intelligence to support the teaching and learning process, but also on the way that Artificial Intelligence systems in the workplace impact on education systems.
I am also Director of something called ‘Educate’, which is a programme funded by the European Union to support relationships between start-ups and SMEs, many of whom are using Artificial Intelligence to build educational technology applications. I hold the Francqui Chair at the University of Leuven. I am also a Trustee of the ‘UFI charitable trust’, which funds projects looking at post-16 education, and I advise Cambridge University Press and McGraw-Hill. I am also part of ‘Partnership on AI’, chairing the AI and Society SIG (Special Interest Group). And I’m president-elect of the Artificial Intelligence and Education International Society.

In terms of policy work, I have provided evidence to House of Commons select-committee in the UK on Artificial Intelligence and Robotics, the House of Lords select-committee on Artificial Intelligence and I am a member of the Topol Review on AI and Medicine.

Ann Mettler: Thank you very much. And finally, Professor Steels.

Luc Steels: Yes, thank you so much for inviting me here. My name is Luc Steels, I’m from Belgium and I have been working for more than 40 years on AI.

I was first at the University of Antwerp studying languages and literature, going into computational linguistics. But then, realising that computers were important, I went to MIT and I studied computer sciences, and specifically, Artificial Intelligence with my adviser Marvin Minsky. Then I worked in the US with Schlumberger, which is a geophysical measurement company, and came back to Brussels in 1983. I founded the AI Lab at the Free University of Brussels (VUB) in 1983, which became quickly one of the biggest AI labs in Europe. And then I was a Professor at the VUB, abd Chairman of the department for a while. In 1996, I was founding director of the Sony Computer Science Lab in Paris, and more recently I have been working as an ICREA research professor at the Institute for Evolutionary Biology (UPF-CSIC) in Barcelona.

So I have worked in many areas of AI: computational linguistics, knowledge-based AI, language evolution, and also on ethical issues. I have participated in the EU Framework Programmes [for Research and Development] since 1982, in fact, so I was already there in the pre-ESPRIT (European Strategic Programme on Research in Information Technology] era.

Ann Mettler: Thank you so much. I’ll now hand over to my colleague Mario Mariniello.

Question 2: What is Artificial Intelligence and which trends are shaping it?

Mario Mariniello: So we now move to question two: ‘In your own words, what is Artificial Intelligence? Please provide the definition that you would consider the most accurate. Please also share your general views on global trends and, in particular, on the current and future prospect of AI development in Europe.’ For this question, you have three minutes each and we start with Professor Crootof.

Rebecca Crootof: When I was growing up, I was told that Artificial Intelligence would exist when a machine system could beat a human being in chess. And then that happened and it was dismissed as mere programming. And since then we’ve had Watson win at Jeopardy, AlphaGo beat the human Go champion, and these feats have been admired but also minimised as weak AI – just Artificial Intelligence that was limited to only accomplishing a single goal.

So I want to point out that as what was recently unimaginable becomes increasingly commonplace, the goalposts for Artificial Intelligence keep moving. Today’s Artificial Intelligence is tomorrow’s simple software programming. In light of the ever-changing nature of what constitutes AI in the popular consciousness, if not in technical definitions, I think we should avoid general definitions that attempt to draw lines in the sand. Instead, we need to recognise that AI is a spectrum. It’s a collection of methods and techniques for making computers more likely to achieve a goal. Even more importantly, we need to be thinking about what goals and parameters we set for these AI systems and considering the social, ethical, political, strategic, and other implications of granting increasingly independent systems with decision-making capabilities power over our lives.
This is a particularly productive time to be considering these issues, as relatively recent technological advances – increased capacity of computing power, the emergence of global digital networks, advances in distributed computing, and access to extensive training data – have resulted in relatively recent dramatic breakthroughs in machine learning (the branch of AI that explores how machines can improve their performance through experience). So, as a result, weaker forms of AI are already integrated into many aspects of our lives, providing answers to our search queries, auto-correcting our misspellings, filtering out spam emails, improving diagnostic imaging, assisting us with driving, peppering us with targeted advertisements, suggesting bail amounts and sentencing times, and allowing for the development of increasingly autonomous weapons systems.

**Mario Mariniello:** Thanks. Over to Professor Dignum.

**Virginia Dignum:** In my own words, Artificial Intelligence is a discipline which studies and develops operational artefacts that exhibit properties of autonomy, interoperability or interaction and are able to learn from those interactions. AI is not a noun and AI is not an adjective. It’s a discipline. In this discipline, it is increasingly important to consider the societal, ethical, legal and cultural impact on the world around us. As such, I see it as being developing systems that are complimentary to human intelligence and complimentary to human activity and not merely as replacing humans on those activities.

Concerning these recent developments and the direction that I see AI going at the moment, I must say I’m growingly concerned about how the development is progressing. I have been doing AI since the 80s. I have passed several of the winters of AI and I’m wondering if we are not moving into the next winter. Why do I say that? It’s because we are too much concerned with the very low hanging fruit applications of extensive computational power and the big data availability and we are only looking at the current state of the art and looking at those low hanging fruits. All the complexity – the problems which exhibit more complexity and further thought – we are, at this moment, not looking at. So I wonder how we can move in that way. Additionally, Europe has quite a lot of capability and has been traditionally very much able to generate create all kinds of technologies.

So I think that, actually, **Europe is very well placed to move Artificial Intelligence into the next generation and take into account other types of technologies and methods rather than correlation-based methods that we are looking at the moment.** I see a very positive future for Europe in the development of the next step in AI.

**Mario Mariniello:** Professor Fung, please.

**Pascale Fung:** I will provide a more technical definition of AI. I think Artificial Intelligence systems have to have a certain amount of machine learning. Without learning, it’s a mere pre-programmed expert system. A modern Artificial Intelligence system manifests a certain amount of autonomy and also learning.

Having said that, I would say that there were, in the past, two schools of AI research: one based on the Marvin Minsky tradition and the other coming from the Shannon school of information theory. In the last few decades, we have actually been using AI in the field without really referring to it as AI. For example, a large number of call centres have been automated by interactive systems and speech recognition since the 1990s. However people were not really referring to them as a AI even though they are. Online machine translation has also been around for almost ten years: people have been using Google Translate for almost 10 years, and that’s based on statistical modelling first and more recently on neural machine translation.

So, we have been living with AI for quite some time without general panic. And I think recently there has been a surge in terms of AI research breakthroughs that came from the breakthroughs in neural networks, which is also a technology that has been around for 30 years. And this breakthrough of research results comes from actually a perfect storm of availability of big data, availability of very powerful GPU [graphics processing unit] machines, which allowed researchers to invent and experiment with neural networks. That was not possible in the previous decade, so this perfect storm created a marketplace for applications.
The speed of application of AI technology from research has just gone really rocket high. And actually I do not see an AI winter this time around, compared to previous times, because this time, everything is quite real and there are real research breakthroughs. It’s not just hype, and real applications coming from industries such as the financial services industry and the medical field are very present (precision medicine and all that).

And for Europe, I see the challenge as being a competition against North America and China. A challenge coming from the lack of, perhaps, a single linguistic market, which is very important for big data. But Europe has a tradition of very sophisticated STEM [science, technology, engineering, and mathematics] education – especially mathematics-based – and that is the advantage of future of research in Europe.

Mario Mariniello: Thank you. Professor Luckin.

Rose Luckin: I find it useful to adopt a less techno-centric definition of Artificial Intelligence. I see it as being about the study of intelligence in an interdisciplinary way, in looking at philosophy, psychology, linguistics as well as computer science and Artificial Intelligence techniques. I think that the growing propensity to look at Artificial Intelligence purely from the technology point of view runs into the ground when it comes to the dialectic between human and artificial intelligence because that relationship between human and artificial intelligence is actually what we’re really concerned about.

Originally, Artificial Intelligence was also about trying to understand human intelligence by building systems to model some of the behaviours. Now these days we have more evidence from neuroscience that helps us to understand human intelligence, but we still don’t have a full grasp of human intelligence. So we need to bear that in mind.

I think context and culture are hugely compounding factors to the way in which Artificial Intelligence will move through society. And of course it’s up to us as humans to make the decisions about what we will and will not allow. I think there’s a risk that we overvalue the type of intelligence that we’ve been able to build in artificially intelligent systems. So yes, we can see those systems as technologies with the ability to perform tasks that would otherwise require human intelligence such as visual perception or machine translation, but human intelligence is much richer than that. And there’s so much of meta-subjective, meta-cognitive, meta-emotional, meta-contextual human personal self-efficacy that is way beyond the realms of any Artificial Intelligence systems that exist at the moment. And we need to bear that in mind because if we define Artificial Intelligence too closely aligned to human intelligence we run the risk of undervaluing human intelligence and overvaluing what we’re getting from Artificial Intelligence.

So the trends are certainly being driven by the perfect storm, absolutely. Big data, processing power and fine algorithms, and deep learning can do many, many useful things. But I think if we don’t take notice of these really big contextual and cultural factors that will impact on how much those AI technologies achieve their expectations, then we run the risk of – I hate to say another winter – but at least autumn because human expectations will exceed reality.

Mario Mariniello: Thank you. Professor Steels.

Luc Steels: Thank you. I would like to amplify what my colleague here (Rose Luckin) just said, namely that there are two aspects to AI. One is scientific: AI is a field studying intelligence and, more broadly, activities of the mind, of being in the world and so on. This is very important because there are many aspects of intelligence that are not yet covered, nor even studied today. So if we just look at it from the technological side, we are missing where the field could be going.

As a scientific field, AI is very different from psychology and from neuroscience because those are observational fields, whereas here you do experiments, so you build systems in order to find out how certain aspects of mental activity could work.
Now of course there’s the technological side which is today very important. Thirty years ago it was not so important, but now it is. And for that, the goal of AI is to provide a set of methods and techniques. It is not an object of some sort that you can add to a system like that. It is a set of methods, and you need to apply them; they don’t come for free. These methods and techniques are intended, first of all, to make machines more intelligent so that they are better decision-makers, but also, possibly, more autonomous.

With respect to intelligence, it is important to realise there are several approaches and I guess the two most important ones are the knowledge-based approach – where you start from human knowledge, human concerns, and try to acquire that knowledge and translate it into a system. And then the data-based approach that starts from data, which is the machine-learning approach.

And, just quickly, with respect to winters: I think there have already been two winters and there will be others. It’s the cycle of things. If you look at other fields like human genomics, it’s also cyclical. Ten years ago we were promised a human genome that would end all diseases. Today, we don’t hear very much about it anymore. So there are cycles in every field and there are also cycles in AI.

The main things we have to worry about are: overestimation – with the consequence that unwarranted fear holds everybody back – and underestimation – with the effect that we don’t seize the opportunities that are already there with AI. We need realistic assessments and stable infrastructure.

Mario Mariniello: Thank you. Finally, Professor Crawford.

Kate Crawford: So you’ve heard some very different definitions today and I would suggest that that’s actually a part of the problem with Artificial Intelligence. The term is simply so big. It has been used since 1956 at the Dartmouth Conference, and it has changed with every decade. So what I’m going to do is to build on Professor Luckin’s comments and to suggest that Artificial Intelligence actually has three components: the technical, the social and the industrial.

First, when people say AI today, they generally are referring to a constellation of techniques from natural language processing to facial recognition to deep learning. And these technical capabilities have become a great deal more powerful in just the last eight years. As Professor Fung noted, there has been a perfect storm of increases in computational power, better algorithms and vast amounts of personal data that have primarily been collected by a small handful of tech companies. AI systems allow you to detect and capitalise on patterns from the intimate to the interstellar, at scales that can be very small – it could be at the level of a cell or a quark – to the size of entire societies or ecosystems. These multi-scalar techniques allow us to see connections in data sets and draw conclusions that can be incredibly powerful. This creates structural advantages for many industries, including insurance, energy, finance and politics.

So AI systems build a model of the world, in the machine learning era, by being trained on vast amounts of data and then they can make predictions and determinations from that model. Now this training data of course, reflects the history and context where it was created. What that means is that, despite our hopes that AI might avoid or overcome human limitations – be that bias, discrimination or exploitation – we’re starting to see those issues built into the AI systems of today.

Second, AI is not just technical approaches, it’s also social practices: how humans are being classified into categories, the way that data is being gathered and valued, how society is understood and which problems we are prioritising. These factors are powerfully shaping the types of AI systems that we have today. And these social practices are also being led by a field which is very demographically homogeneous in terms of socio-economic backgrounds, in terms of gender and in terms of race. This also affects the kinds of problems that AI is choosing to address and which populations are best served by these tools.
Finally, AI is also an industry. The current ascent of AI is being driven by the consolidation of the tech industry through the last 10 years with social media platform growth to see the global powerhouses that we have today. Now currently, there are really just a handful of companies who are able to do large-scale AI because it simply requires so much recent data and extensive infrastructure and a high-end skilled workforce. These factors have been concentrating power over the design of AI systems into relatively few spaces.

Geopolitically, the AI industry is concentrated in just a handful of countries. We can study the number of AI patents to see that the concentration is really in the US and China, with some activity in the UK and in parts of Europe and Japan. And this is happening at a time of rising inequality and continuing geopolitical tensions that I think are creating significant new power imbalances. The global North are becoming the AI have, the global South are becoming the AI have-nots. And I think how we choose to deal with this imbalance, particularly in Europe, is going to be a crucial public policy issue for the 21st century.

Question 3: Benefits of Artificial Intelligence and barriers to its development

Ann Mettler: Thank you so much. We will now continue with question three. Please elaborate on the potential benefits yielded by Artificial Intelligence to society: citizens, business and public sector. You are welcome to describe concrete AI applications to illustrate your point. In addition, please elaborate on barriers of any nature that could hamper the adoption and development of Artificial Intelligence. And this time we start with Professor Dignum, please.

Virginia Dignum: Thank you. As we saw from the answers to the previous question, AI is not just one thing. There are many types and it comes in many sizes and there are many differences in the types of applications but also the types of philosophies behind it. I believe that, ultimately, taken as a whole, AI will have the ability to make machines more useful, which will enable us to create benefit for us, for people, for our society and for the sustainable well-being of all of the planet.

However, we are dealing with issues of how to integrate these systems and possibilities in our society. And that will only be successful if we can and are able to align our development of AI systems with human values and human principles. AI – and all the systems built with AI – are tools, artefacts, things that we build, and therefore we are responsible for those systems. We need to take accountability and transparency into those systems in order that we ourselves, our societies, the users, the governments, can take responsibility for the use – or not – of those systems, and also for results and the possibilities that those systems will bring us.

Contrary to the opinion of several of my colleagues – not necessarily around this table – I believe that regulation will contribute and be a necessary step for building better AI systems. And the example that I would like to give is an example from the automotive industry: when we introduced regulation that required cars to have catalysts, we have brought the car industry into developing better cars. So I do believe that we need catalysts for AI systems in order to ensure that they align with our values.

One of the big issues here is that when I said ‘our values’, it is something which is of course very difficult to understand and to agree or decide on. What are the common values of the users of AI systems, especially taking the universal possibilities of AI? And there is where the public sector – and especially Europe at its level of governance – has a very important role to play in order to ensure that those values which we can define as European values are taken seriously and are central to the development of AI systems. We should not first develop the systems and then we go with the checklist to see if they fulfil those values. We really have to turn it around and determine which are the values, the principles that we want to see at the core of these systems and take those ones as the basis for the development of systems.

Ann Mettler: Thank you so much. Professor Fung is next, please.
**Pascale Fung:** The areas in which I think AI can benefit society almost immediately include precision medicine: AI will be able to enhance the abilities of human doctors for better diagnosis and better pharmaceutical products. But final decision-making should remain in the hands of certified medical practitioners.

Another area where AI can improve efficiency and lower costs is manufacturing. And that will actually benefit a lot of the so-called global South and sort of bridge the gap between the AI haves and the AI have-nots.

And the third area is the immediate benefit to be gained from AI in the financial services industry. And you might wonder why that is beneficial to society. Well, if AI is applied to ensure ethical practices of financial services, to enforce compliance in the financial sector, it can prevent market crashes and market manipulation and that will ultimately benefit the society.

I want to also address the issue of what might hamper the development of AI. I think one thing is actually cultural in the sense that today we see the application of AI in major markets such as North America and China, which have very different ethical standards and principles both from each other and from Europe. It would be very misleading to think that we could all force the ethical principles of one culture to another. Therefore, I think we need a global collaborative dialogue – if not collaboration – on AI guidelines (we might not want to call it ethical principles, because those are very differently defined in different cultures) and, more concretely, on some kind of trading guidelines, in the sense of ‘how can we ensure AI safety, AI security, AI fairness?’, very much like similar strategies and guidelines for trade between different markets.

So, I regard AI as a very practical trading – not commodity, but – necessity, for different countries to work together. And I do see that starting to happen. There are various AI guidelines and organisations being formed in North America, in Asia, in China. And this is very new. People are talking about it. Different nations are looking at this as being very important, and I think it’s very important that we all work together.

**Ann Mettler:** Thank you so much. Professor Luckin is next, please.

**Rose Luckin:** The biggest benefits from Artificial Intelligence will come or do come from when we look at the problems that we need to solve. So, taking problems, unpacking them and looking at how we blend artificial and human intelligence together correctly.

On medicine, I completely agree, there is huge potential for profound impact. But, if we take something like IBM Watson Oncology – yes, very efficient; can analyse lots and lots of information about cases, about treatments – but when it comes to that conversation with patients, we need our human clinician to have that difficult conversation with the patient, to explain what it means for them. So it’s when we get the blend right that we will make the biggest impact.

But in terms of impacts that we feel already, it makes my life so much easier when I travel in a city where I can use Citymapper and the data has been opened up and I can find my way around – taxi strike or no taxi strike. In terms of business, it can drive much more efficient marketing, for example targeting messages, targeting customers. For the public sector, again, use of AI could really improve efficiencies within the public sector.

But, again, it comes down to being interdisciplinary. If I look at medicine, some of the really innovative developments around swallowable AI mean that we have to look at nanotechnology as well as AI. So it’s about blending different technologies together as well as different disciplines to understand what’s going on in terms of the medical benefits of the device.
But even where there are these tremendous potential benefits in terms of efficiency, clinical decisions, diagnosis, treatment planning, imaging, administration within the health service alone, there are resourcing issues and a lack of readiness to take advantage of what AI might bring. And in terms of the increased productivity that is often heralded from using AI, for example in manufacturing but also in other parts of the business world, that also would depend on how well trained the human members of staff are to work alongside the AI. Because increasingly, we’re looking at more augmented situations, which means that, to actually make it work, we have to make sure that humans are educated and trained to understand how to work alongside the Artificial Intelligence.

I think AI can really drive innovation by helping us to look at things differently, from analysing large data sets and seeing things in a way we haven’t been able to see before. And there’s some really nice examples of improved efficiencies coming through innovations such as through Google Kaggle competitions where people can compete to build the best version of something to improve a particular public sector problem. So again very problem-focused.

The benefits will also come through disruption. But that’s a double-edged sword. Disruption can be a negative as well as a positive: In some areas disruption is a good thing because it makes us throw things up in the air, look at them differently and think ‘oh, actually we could do this differently now that we have these technologies to help us’. But again, there’ll be social and economic factors as well as the technical factors that will impact on the extent to which that disruption is positive or negative.

In order to accelerate adoption, we need to make sure that, as a human workforce, we’re ready to use AI effectively. And if we do that properly we can really help to democratise a lot of opportunities for people. If I look at the area of education and special education for example, the existence of an audio interface is amazing for people who can’t operate a keyboard. So there are lots of potential ways in which we can benefit in very special ways if we make sure that that kind of AI is promoted, because often these are communities who don’t have a very strong voice. So we need to make sure that we look at the communities who could really benefit from AI in this way.

And finally, I think in order to really assure ourselves of the best benefit from Artificial Intelligence, we need to absolutely demand intelligibility. Whether that is through transparency or ‘explainability’ – we have to build intelligible Artificial Intelligence systems.

Ann Mettler: Thank you so much. Professor Steels.

Luc Steels: Yes. Of course, I also believe that tremendous benefits are possible. One is obviously for companies and I think it’s a matter of economic survival for European companies to engage profoundly with AI. It gives a competitive advantage, for example by linking producers and consumers, or for identifying, designing, or manufacturing products. It’s about keeping jobs, creating jobs. And I think it’s imperative to realise that.

But I want to focus more on the public sector, which is not so much discussed usually. By that, I mean what the citizens can get out of AI. And I think there’s actually a lot but of course we have to make it happen. It won’t happen by itself. There are already projects where citizens are empowered because they can use IT. We should do the same with AI.

One example that I was involved in is a system called Noise Tube, which enabled citizens to record, measure and upload information about noise in their environment. This then provides a way to communicate with municipalities about noise problems that exist in certain areas. In Europe we place a lot of emphasis on social structures and so there is a lot of demand for it but also a lot of developments. So that’s another area for improvement with AI, so that people can really see a benefit for themselves. It’s not about efficiency. If that is the only thing we focus on, we would lose. Instead, we have to really look at what can make life better for citizens, for users, and those working in administrations.
There are two other application areas which I will just briefly mention: In Europe, we place a high value on culture and we also have tremendous historical resources. Museums, musical catalogues, and things like that. So to have access to this and also to get a historical perspective, which is today not the most obvious thing to have, I think that AI is relevant. It could help to bring history to life and to refocus society on culture as one of the important things that enrich people’s lives.

And finally, I would say that education is another important area where, again in Europe, we have a lot of strengths, we have enormous traditions. All European countries place a lot of emphasis on education. There is a sort of general consensus that education is very important and should be accessible to all. And here, again, there are opportunities. For example, it doesn’t only have to be mathematics or science, but also music education. There have been great projects funded by the European Commission on building MOOCs [Massive Open Online Courses] or tools so that everybody can get a good musical education. Also for language tuition, there are opportunities of course.

So I think these are potential areas that do not fall within the more narrow, very profit-oriented, efficiency-oriented domains – those of the sort: ‘we’re going to be able to scrap so many jobs’ and things like that. We should switch to really thinking about benefits – human benefits for as many people as possible.

Ann Mettler: Thank you so much. Professor Crawford is next, please.

Kate Crawford: It’s wonderful to hear so many benefits articulated so well here today. I would certainly agree that where I’ve been the most impressed by AI has been in the areas of disease detection, energy efficiency, anomaly detection in radiology, speech translation, environmental management and climatology.

Some of you might have just seen the new article that came out in Nature just a week ago that indicated that AI was used to discover six thousand new viruses (that was through two studies). Truly an extraordinary discovery. And meanwhile, across the other side of the Atlantic, the UK National Grid is now trying to use AI to manage electricity consumption and essentially they believe that they can reduce energy consumption by 10 percent, which is significant. We’ve also seen some excellent examples in how AI can be used to track endangered species – something that is traditionally very expensive and difficult to do. And of course we know that animal poaching is so severe in parts of the world that some populations are almost completely dying out. And researchers are now using AI systems to track migration patterns to be able to select particular areas where the most effort should be directed. These are just a few of the examples that I think are really interesting in terms of AI’s potential to operate beyond the previous tools that we’ve been working with to essentially reveal new forms of knowledge.

I’d also direct you to what I call the FATE area of research: fairness, accountability, transparency and ethics. This area has been getting a lot of attention in the US just over the last few years and it’s now also growing in Europe. And this also connects to this idea we have discussed today of the interpretability of AI. Trying to make AI explainable is something that has been very much within the FATE area of research, so that’s something that I think is particularly important here.

But I would like to also remind us that when we discuss benefits, we should be looking for measurable evidence and not rely on marketing assertions. For example, we’ve already seen cases where excessive claims about the benefits of AI have really backfired. IBM Watson’s partnership with MD Anderson, which is a major cancer research hospital in the US, collapsed after only three years, after it was shown that Watson had not made improvements on the health outcomes of their cancer patients. And with the recent death of a pedestrian by an autonomous car that was operated by Uber in Arizona, I think it’s a reminder that our excitement should not be allowed to overtake much needed research and development into safety and potential harms.
So I would suggest that when we hear ‘benefits’ we should often think about flipping it to see the two sides of the coin. When we have advances in radiology, does this also result in a shift of power from public health systems into private hands? Do improvements in manufacturing bring all the spoils to a few incumbents, or does it really benefit the global South? Even AI for tracking endangered animals, could that then be used and gamed by poachers to try and find precisely those same animals? So I think we can see a lot of double sidedness to these sorts of examples and I think it helps the analysis to hold both in our minds at once.

And finally, we have to recognise the importance of trust here. Because the benefits of AI will only be trusted if the claims match the reality and if people are not excessively exposed to risks. I think the recent revelations about data analytics and behavioural targeting seeking to influence election outcomes are one of these flip sides of the tech optimism about the democratic effects of social media and data sharing that we saw just eight years ago. So when we talk about the benefits of AI, I think we need to ask: ‘Who are the actors deploying these technologies and to what ends? Who benefits the most? Who bears the most risk and what are the unintended consequences and the possible impacts on the public, particularly the most vulnerable or marginalised groups in society?’ Thank you.

Ann Mettler: Thank you so much. And finally, Professor Crootof, please.

Rebecca Crootof: Thank you. As someone accustomed to focusing and thinking about the harms associated with Artificial Intelligence, I really appreciate the opportunity to think about the benefits, to recognise that Artificial Intelligence systems are tools. And, like all tools, they have the potential to expand human capabilities. I usually focus on how they extend them in negative ways, but obviously – as evidenced by many of my colleagues’ comments – there are many ways that they can be extended in positive ways.

I do want to emphasise that AI systems attempt to optimise some behaviours or achieve some goal, and it is up to human beings to determine what that goal is and what the parameters are on the process for achieving it. And it’s up to human beings to consider all of the consequences associated with achieving that goal, and the implications of it. I want to also emphasise some of the most beneficial aspects of AI might come from human-AI teaming. There’s the example of AI systems that can beat human players at chess. But AI-human systems together working collectively are better than either individually. Wikimedia editors, for example, are assisted by bots to help create a global source of information that wouldn’t be able to exist and be curated otherwise. AI can assist in content moderation – not by automatically censoring content, but by identifying posts with problematic language and asking the poster if they want to reconsider it, so they can provide a nudge without necessarily being used as a heavy hammer or censor. As other people mentioned, AI can assist with medical diagnostics, with legal discovery, with urban planning, with space exploration, and with addressing the impacts of climate change.

All in all, we should be considering the myriad ways in which AI can be of assistance to human beings and leverage the respective human and machine intelligence. I am going to reserve the rest of my time to talk about harms.

**Question 4: Harms and ethical dilemmas of Artificial Intelligence**

Mario Mariniello: All right. So we move now to question 4: Please elaborate on the potential harm that could be caused by Artificial Intelligence. Please state what worries you the most when considering the impact that Artificial Intelligence may have on human beings, identify the sources of highest concern for society and describe the ethical dilemmas that you consider most pressing. And again, you have five minutes each. We start with Professor Fung.

Pascale Fung: From my near 30-year career in the field of speech processing and natural language processing and, more recently, affect recognition, I think what worries me most is that most of us engineers and scientists working in this field have never given a thought to the impact – both benefit or negative – of what we do on the rest of the society.
I remember – and it’s still the case – most of our research and our graduate students, and schools are funded through defence projects. My very first project was to provide speech commands to fighter jet pilots to launch bombs. And we have been telling ourselves ‘our science is quite independent from politics and society’. I think more recently we have seen that this is not the case at all, and that people can use our technology to create fake news. They can create fake IDs, fake humans, imposters, to manipulate both the market and political campaigns, to undermine the democratic system and the financial market, and so on and so forth.

I have become very concerned about this in the sense that our AI workers, our creators of AI technologies, our engineers today and in the future are still not educated at all about ethical principles and the ramifications of what we do. **We do not have the kind of Hippocratic Oath that medical professionals have to take.** Recently we’ve had a couple of, I would say, voluntary declarations from some of us, to vow that we will not do harm by creating technology that would negatively impact the society. But I think that is not enough. **So my greatest concern really is AI education and education of the creators of such technology.**

Because AI is still human-centric, it is very, very important for us to educate our engineers and our scientists in terms of the ramifications, the philosophy, the ethical principles and standards. This means that our education system has to change and reform dramatically. Because traditionally, STEM and humanities education have been quite separated and we cannot afford that. So that’s my biggest concern today.

I have other concerns about the impact of this technology but the source of such concern really comes from the education.

**Mario Mariniello:** Thank you very much. Professor Luckin, please.

**Rose Luckin:** I think it’s very tempting when we think about harm from AI to contemplate the possibility of rogue systems taking over the world. That’s not something I adhere to but what I would say is that harm from AI systems is human-driven: we design the AI systems, we train them with the data sets that we train them with. If we do harm with our AI systems, then we have to take responsibility for that.

And I think there’s a huge responsibility on society to try our best to put in place frameworks and definitely education programmes to ensure that we minimise the risk of that. But I don’t think anybody should run away with the idea that it’s not going to happen because it is, because unfortunately that is human nature. So it’s really for me about what we do most to try and prevent that kind of harm.

**The sorts of harm that I worry about are very much connected to risks for underprivileged communities who have less access to education about Artificial Intelligence opportunities to influence the training data that AI systems are built on.** I worry about that and I think we need to do something to make sure that diversity is increased.

I also worry that as humans we’re inherently lazy. And so if a company comes along and says ‘hey you don’t need to do that any more because we’ve got a system that can do that for you’, that’s very tempting. And to an extent that’s a good thing, possibly, but to an extent we may offload the wrong things to Artificial Intelligence systems if it dumbs us intellectually. I mean, these days for shopping, do I need to make any decisions myself? I’m offered a whole plethora of recommendations for books, for movies, for food. There’s a risk that we lose the ability to make good decisions and this taps into what you were just saying, Pascale, about fake news and worries around that.

A lot of that is driven by problems within our education system that are not just about the lack of STEM education – though I absolutely acknowledge that’s extremely important – but it’s about not helping people to ‘know what evidence really is’, how you justify making a decision, how you justify whether you believe something or not.
We’ve tended to move towards a system that proposes that students are exposed to knowledge. But they have quite a poor, quite naive personal epistemology about where knowledge comes from or what evidence is, which leaves them prone to be sucked in by fake evidence and less able to make critical judgements about it. And that is where there’s potentially a lot of harm. Yes, of course, we can expect supercharged cyberattacks so that means that we have to anticipate those happening – and again it would probably be the most vulnerable who would suffer the most from those – so we need to make sure that we’ve got good regulatory processes in place. But we also have to make sure that people understand what AI is, what AI can do and what it can’t do. We must not dumb the population down by saying ‘it’s too hard, they can’t understand it’. We must help people to understand what a system can do with their data and with its algorithms, whether they benefit from this system, and what the less desirable consequences might be. So it’s a real job for the education sector not just to educate those who build AI but also educate people to understand enough about it.

And of course there is the question of ethics, which I know you’re looking at a lot so I won’t dwell on it because I know you have a lot of information. But obviously we need to have an extremely strong ethical framework in order to try and make sure that the systems that are developed are for benefit rather than for harm. And I will stop there because I’d like to save time for the last one.

Mario Mariniello: Thank you very much.

Luc Steels: First of all, I do not believe at all in this idea of a super-intelligence that will enter this room and do all sorts of bad things to us. This is science fiction. This is Hollywood, this is not a reality. And comparing the risk of AI to the risk of nuclear war, as some people are doing, I find a total exaggeration which is extremely unhelpful. Tomorrow in ‘DIE ZEIT’ (it’s a German newspaper), there will again be a whole article with this kind of talk.

Now this doesn’t mean that there’s no harm. And I think we have to say that the risk is mostly in the misuse – potential misuse or real misuse – of AI by politicians or business people, who might employ 20 year-old hackers to do something. We have seen this in the Cambridge Analytica example recently. You could say that it is the developer who implemented the system who is responsible, but actually it’s the people behind all this which we have to question. We can give as many ethics courses as we want to our AI students. But if they get caught in this kind of web of manipulation it will be very difficult to resist.

There is also real harm possible. First of all by overestimation, because there can be inappropriate applications of the technology which can be very insensitive. If you talk about AI-based decisions on granting parole, for example, or getting social benefits or insurance, these are all things which affect people. We have to be pretty sure of what we’re doing. And I believe in most cases it is foolish to assume that we just can leave decisions to an autonomous AI system.

There’s also this belief, due to overestimation, that humans can be replaced. Now this is, I think, the most stupid idea we can have. You know, if you have an insurance company, you build an AI system based on what those people as experts were doing and at the end of the day you say ‘Ok, goodbye everybody, we now have our AI system?’ I mean, that is totally stupid because knowledge is something that’s evolving all the time and it’s related to context and social issues and all that, which are not captured by having a machine learning algorithm run over the data.

So another issue which you raised also is that this overestimation can lead to our own skills and our own knowledge not being further developed, which would be another great mistake. I mean, most people can no longer read maps, you know, because why would you learn to read maps, given that you have a smartphone to guide you? Why would you learn how to write, because you speak and then the computer writes it for you? I heard some parents say ‘well, we don’t need to teach our children how to programme because AI is already writing all the programmes we will ever need’. And so on. This is something that we have to worry about. Such overestimation could lead to the collapse of all sorts of systems because the knowledge will no longer be there to maintain it, let alone to develop it further.
Now the next thing I would like to say is this power asymmetry which may come from some people having access to AI and developing it, while others don’t. We risk that those who have it can do all sorts of things like hidden manipulation of citizens in democratic elections or market manipulation. Moreover, the hunger for data is such that there’s a disregard for privacy, a disregard for ownership and all those kinds of things. We have to really worry about that attitude to power which is part of human nature. We have to seek that it is balanced in some way. But this is not just an AI problem; it’s an ethical problem at large.

Mario Mariniello: Thank you. Professor Crawford.

Kate Crawford: When I think of examples of greatest concern that keep me up at night there are essentially three categories: social control and asymmetries of power; the magnification of discrimination and inequality; and the erosion of significant rights and liberties.

Let’s take, for example, China’s emerging credit score system. So this is essentially a citizen ranking and blacklisting system, like a meta credit score on essentially everything that you do. It can track what you buy, your geolocation patterns, your social graph. Your scores can go down if you turn up to work late or if you play too many video games or leave negative product reviews. And they can go up if you say positive things about the government or you pay your bills on time. The score can also determine many things like what sort of job you can get, whether you get a mortgage, and – as we discovered in a story that was just published two days ago – whether you can travel. It has now been shown that 12 million Chinese citizens have been blocked from domestic travel.

So while some people are saying that this could only happen in a society with extremely centralised political power like China, I would suggest that we’re already starting to see concerning signs in the West. Social media companies are already patenting the ability to offer loans based on the credit scores of your friends; health insurance companies are modulating premiums based on Fitbit data and other wellness data; and HR companies using facial recognition in hiring, to assess the micro-movements in the faces of job applicants and then hiring them if they match the face profiles of their highest performers. This is already being deployed.

So it is concerning to me not just from the perspective of privacy and the centralisation of power, but also from the aspect of discrimination. We’ve seen how AI systems trained on historical data can produce biased results. In the US, we’ve got racial discrimination in risk assessment scores used in post-trial sentencing in the criminal justice system. We’ve seen gender discrimination in search results and we’ve seen age discrimination in targeted job ads. So I think we need to be extremely careful to ensure that these forms of structural discrimination and historical inequality are not being reinforced by AI systems.

There’s also some concerning developments that I’m seeing in AI research: back in 2016, two researchers published a paper where they claimed that they created an automated criminality detection system just by looking at a photo of somebody’s face. Now, in fact, I think the underlying claims of this study are deeply suspect and ultimately similar to pseudo-scientific approaches like phenology and physiognomy in the 19th century. But unfortunately, we’re starting to see a rerun of those same ideas in the AI studies of today.

Now, again, you might have seen a study that came out just two months ago by Wang and Kosinski that claimed that it could detect sexual orientation from people’s faces. They did this by scraping dating websites and Facebook and they claimed high levels of accuracy in identifying gay men and women. Now many of us, including some of my colleagues at this table, expressed real concerns about the sample size and techniques in the study. But I think beyond these methodological critiques, we really need to think about the ethics of AI classification. Particularly in this case – given that homosexuality is still criminalised in 78 countries, some with a death penalty – I think it’s a very strong reminder that our responsibility in how we make AI tools has never been greater. So we need to think about how AI tools connect with existing forms of power.
Obviously in the last week we’ve seen a big example of how large scale systems can seek to manipulate news and opinion in ways that are very hard to detect. I think the Facebook-Cambridge Analytica story is useful to me as a reminder not just of how tools can be abused – which is something we’ve talked about here today – but of the long-term implications of business models that are based on maximally exploiting people’s data while only giving them marginal control. So *I think we need to urgently strengthen fundamental research on the social implications of these technologies and in particular to really emphasise the importance of interdisciplinary research.*

You’ve heard it from several of my colleagues today that we need to rethink how computer science and engineering is taught. Right now, we have computer scientists who are being asked to design systems for law, for hospitals and for schools. And this is simply outside of their existing training. So what I’m most excited about to address some of these threats is that the AI field will grow to encompass many disciplines: law, anthropology, sociology, science and technology studies and history, among others. And indeed this is why we created the AI Now Institute, because I think *it is urgent that we start to broaden this definition of what the AI field is, to move beyond the solely technical, to assess what are essentially socio-technical concerns.*

**Mario Mariniello:** Thank you. Professor Crootof.

**Rebecca Crootof:** Like other tools, Artificial Intelligence systems have the potential to increase human capabilities. Accordingly, AI systems permit a variety of intentional and unintentional harms. I am only going to discuss three of the many categories today: actively malicious action, unavoidable accidents, and issues arising from humans trusting AI too much.

So first, as detailed in a report that I and numerous co-authors put out recently, *AI enables actively malicious action: criminal hackers can use AI to automate social engineering attacks, for spear fishing, or to detect and exploit vulnerabilities in code.* Companies can use AI to manipulate customers’ actions and behaviour. Governments can use AI to surveil and suppress dissent, to deploy personalised misinformation campaigns and deep fakes that risk destabilising the public sphere. Additionally, there are a number of unresolved vulnerabilities in today’s AI systems that risk malicious tampering by competing companies, hackers, terrorists, and enemy states. These include data poisoning attacks (using training data that causes a machine to make mistakes), adversarial examples, which are inputs that are designed to be misclassified by AI systems (the common example is a stop sign that is read as a 45 mile per hour speed limit sign), and reward hacking, which is an exploitation of flaws in how the AI’s goals are designed. That’s one.

Second, history has shown over and over again that tragedies occur when complex systems managing hazardous material or sensitive information fail due to unanticipated interactions within the system. The Three Mile Island accident, the Air France Flight 447 crash, and the trillion dollar stock market flash crash, were all normal accidents: failures that occurred despite careful design, planning and the attempt to include multiple technical redundancies. *The more power we delegate to complex AI systems, the higher the risk that these normal accidents will occur and will result in catastrophe.*

Third, I think *some of the greatest risk from AI will come from human beings trusting it too much.* This will often take the form of assuming that AI has common sense. But AI is only as good as its data set, and data sets can be biased, incomplete, inaccurate, or inapplicable. I recently saw a joke: a concerned parent asked ‘if all your friends jumped off a bridge, would you too?’ The machine learning algorithm says, very sincerely, ‘of course’. Additionally, AI systems learn to optimise their behaviour to achieve their goal. But they can be simultaneously clever and extremely stupid. So a humanoid robot told to touch its left ear with its right hand might take the most direct route through its head. Or an autonomous vehicle asked to find the fastest route to the airport could jump the sidewalk and mow down pedestrians. Alright, these are exaggerated examples, but they highlight the problem of defining goals without necessarily inputting all parameters.
Petrov arguably prevented World War 3 when he mistrusted the Soviet Union’s early warning system, which had misidentified sunlight reflecting off clouds as an incoming missile strike. One challenge, and an important challenge, for AI developers will be in building appropriate amounts of mistrust for the humans involved in using AI systems.

Mario Mariniello: Thank you. Finally, Professor Dignum.

Virginia Dignum: Thank you. Many different types of harms and issues have been discussed, so I will try to focus on a few that in my opinion are not yet completely discussed. Indeed, AI is an artefact, is a tool that we develop, is something which we as people should and must take full responsibility for.

One of the issues that I see as a harm is from the part of us, the developers, the researchers on AI, that we might be taking too much a too myopic view on Artificial Intelligence and on the capabilities and on the potential of the current capability of it.

Often, we talk about AI and discuss the issue of the algorithm. Algorithm doesn’t mean anything more than a recipe to solve a problem. So that’s what AI systems are: systems which have a very smart, very good recipe for some kinds of problems. Of course, human intelligence is much more than recipes to solve problems. And by focusing and by kind of translating or assuming that that’s all what there is to intelligence, we are losing a lot of other capabilities that humans have and that might also be able to be enhanced and helped by artificial means. And that is what I said at the beginning of my intervention before: the issue of too much reliance on algorithms and on methods that are based on data availability and the necessary computational power for it.

Data hungry approaches have, in my opinion, two issues which we should take into account as we move into other possibilities and other approaches to AI. On one hand, we need all this data to develop algorithms or to train our algorithms and improve all the results of these algorithms. We are also responsible for taking care of this data and – I don’t have to talk about this in this building – and of all the issues of governance and of the institutions that are needed to take care of data.

And on the other hand, there is also the technological aspect of maintaining data. It is a burden on our resources: to store, to keep all this data and to use all these computational algorithms has an impact on our energy resources. Often, data is compared to oil as the new oil. We made a mess with the last oil that we used, so maybe we should take those lessons from history to see how we are going to use it and to deal with big data. And I think that we really need to look at other ways to deal with this issue.

We are taking a too technological perspective on AI. AI is a discipline and is definitely not a technological discipline. It is a discipline which involves many different types of contributions from sociology, from law, from philosophy, from computer science, of course. We have to take that into account and I fully agree with my colleagues that education is one of the issues.

And I think that one of the steps we have to take into the future is this kind of border that we have between STEM education and humanities education on the other side: it is not something which we can hold onto in the future. We need to bring humanities into the technological schools and we need to bring STEM into the humanities tradition.

We really need to take care of public engagement and public participation to ensure that everybody is able to assess and to use the potential. And in that, I think that also the narrative that we are taking on AI is indeed, like some of my colleagues said, kind of detrimental for the good understanding of the kind of capabilities and possibilities that it offers.
Finally, I just want to refer to the issue of democracy and I think – and I will be very quick – I think that democracy has been defined as the least bad of all the governmental systems. And I think that we are going to see a very disruptive impact on democracy – and we are seeing already the negative issues. I think that we really have to consider and look at how we can use AI and the availability of data to make this disruption positive for our democratic process.

**Question 5: Areas of action for public policy and the European Union**

*Ann Mettler:* Thank you so much. We will now turn to question five: Based on your research and professional experience, please identify the main areas of action for public policies aimed to foster the development and adoption of Artificial Intelligence and to mitigate the risk of its potential harm. In your answer, please focus on the European Union: provide your assessment of the current landscape and your suggestions for an effective and coherent European Union Strategy for Artificial Intelligence. You have a maximum of five minutes and we will start with you, Professor Luckin.

*Rose Luckin:* Well, there certainly is an enormous role for public policy in all of this, no question, and it’s both there in terms of strategic direction and regulation and in terms of the education piece. When I read through the list of questions here, they are very interconnected and actually that regulation and education piece is also very much interconnected.

Currently, there’s a lack of coordination and there’s a real need for public policy that coordinates and maintains a common piece across multiple stakeholders. And that’s not there at the moment. We need a Body to steward the governance landscape of Artificial Intelligence at a European level. And we need to ensure that industry and others understand the principles of that overarching structure and that they adhere to them and therefore that means we have to do monitoring. But we mustn’t be reduced to a tick box type process. It needs to be much more meaningful than that. I don’t believe in the idea that it can all be led from companies. It’s great if companies have for example their own ethics boards. I applaud that. Necessary, but not sufficient. It has to be something at a much higher level. Because we must make anybody who is developing or applying Artificial Intelligence accountable for what they do. And we also need to be able to protect the vulnerable, the disadvantaged. You know, that might be people who are stateless, that might be people who are unemployed, it might be people who are just not part of the current conversation about Artificial Intelligence.

We need regulation to ensure that those vulnerable communities are not ignored, and we need to build this framework within which we build trust in the population that they can use AI systems and without too much worry about what’s going to happen. That has to go hand in hand with education. So, for example, it’s absolutely fine to have regulation that means that I have a right to an explanation of a decision that’s been made automatically. But if I don’t understand enough about what that automated decision processing is, I’m none the wiser. And so I’m not really empowered by that. And so we need to give people control over their data. As my colleague said you know there’s this wonderful cliché ‘data is the new oil’ and I always say: ‘Yeah, and it’s unrefined and how we benefit from it depends on how we refine it; but also its wealth is incredibly unevenly spread’. You know, the market in personal data is not benefiting most of the people whose data it is. And therefore we do need regulation in order to ensure that it’s refined appropriately and the benefits are shared throughout society.

So in my view, we need something at a European level to ensure that we exploit AI for the good of society as a whole, not for the good of a small group of large corporates in particular. And that education is a real problem because yes, we need to build AI capacity. **We need to do capacity building in terms of funding research. We need to build systems for companies to help them be assured about intellectual property rights. We need policies in that space. We need to encourage university spin-outs because that’s a great way for research to build innovation. We need to connect people who do research in AI, people who build AI, people who understand the impact of AI on society.**
But one of the real problems we face – and it’s interesting hearing my colleagues already talking about the STEM-humanities divide – is that we have built systems that are incredibly good at learning, incredibly good at learning a lot of the things that we currently teach through our education system. Now we have to work out what we do about that and often people look at me in horror when I say this and say that: ‘Do you mean we shouldn’t teach history or we shouldn’t teach physics?’ No. We have to look at doing it in a different way. It’s not enough to know what the Battle of Hastings was. We need to know ‘how do I know that that evidence is true’. Why should I believe this thing that’s been written about that happened in 1066 in England, for heaven’s sake? How do I know that’s true?

It’s about looking at things differently, so that we educate people to do the things that we cannot automate, not to pit them against the systems that will be faster and more accurate than them. So that regulation and education has to be joined together.

Ann Mettler: Thank you so much. Professor Steels.

Luc Steels: To the first question ‘is the EU strategy needed’, I think there’s no doubt about the positive answer to this question. There are several reasons but it’s clear that if every Member State is, for example, inventing its own rules and regulations, we get chaos. That’s why the European Union was first established: to avoid that, and so that there can be scale-up; so that developments, even in a small country, can be scaled to a larger market, or a larger group of users. So it seems entirely obvious that there’s a role for the EU here.

It has also to do with equality. We have already heard a few times that there’s only certain pockets where AI is currently being practiced in Europe. This is a real problem, because there are a lot of companies left behind. That’s not the idea of the European Union either. So by having a strategy, you can also help Member States that do not have the resources to set up a nation-wide strategy or impulse programmes. But there are very good people sometimes in these countries and then you see them pop up in the Googles and Facebooks of this world, instead of contributing to the local economy.

As for whether an ethical legal charter is necessary, I think I would say yes. Now I’m a little bit worried that all the bad things of this world are now ascribed to AI. You know, the thing we’ve been hearing, everybody is very good at pointing to something bad in the world and then saying ‘AI is responsible for it’. I think that goes too far. After all, the whole scandal that we had recently, just to use that example, was based on developments in the psychology department of Cambridge University, not in an AI department. And a lot of manipulations are based on all sorts of things that have nothing to do with AI. So we should also not go the other way and say everybody can do what they want, but then say the AI people have to be under strict control because they are the bad guys.

In terms of social acceptance, this is a big issue because people read in the newspapers how bad it all is. Soon parents will no longer allow their children to study AI because it is harmful, and neighbours could speak badly about them. I think that’s going too far. The European Commission has an important role to play in improving information on AI. In making clear its limitations, who is doing what, what is the exact role of AI in some developments in society, and particularly also in helping to push applications that are beneficial to citizens.

It’s normal for companies to look for markets in which they can survive or make profits. But there are all these things which are for the benefit of citizens which would not normally be developed for profit motives. And so it’s a matter of stimulation. It doesn’t necessarily have to cost a lot of money. Just amplifying or pointing out possibilities, helping countries or organisations that do not have the resources to nevertheless profit from the potential that is undoubtedly there.

I don’t think we need technical infrastructure. We don’t need to build new hugely expensive mega installations – like in physics for example. But we need propagation of information and amplification of the good things that are happening but are now too often petering out, I should say, because they are not brought into the open enough or they are not brought up in the media that European citizens are using, which are most often not English-speaking media.
Ann Mettler: Thank you so much. Professor Crawford is next, please.

Kate Crawford: So, fortunately, I think there’s an enormous amount that can be done here in the EU both to aid the healthy development of AI but also to guarantee safety and fairness for people here and beyond. So I think rather than being fearful, I think we need to be realistic and to keep our eyes open and to think about four key strategies that I would like to outline for you now:

The first is that core public agencies – and here I’m speaking of the ones that deal with high stakes domains like criminal, justice, health or welfare – should currently avoid using black box algorithmic systems until we have more protections in place. This is certainly something that I’ve done a lot of research on. And to address this problem, I recently co-authored a large scale framework at the AI Now Institute, which outlines Algorithmic Impact Assessments. Now, Algorithmic Impact Assessments are a practical set of steps to account for, test and govern the use of algorithmic and AI systems within core government agencies. Our framework is calling for agencies to begin by just enumerating which systems they are using. That’s certainly more than we currently have to work from and I think it would be incredibly helpful. And then we also suggest mechanisms for public agencies to conduct their own assessments of potential social impacts of the system and then to provide meaningful access to external auditors and to researchers to test how they work, so they can look for these potential harms of forms of discrimination in advance. Ultimately this will also give the public the means to have a say about the use of these systems and their results. That report is going to be published in just a few days.

This connects to the second strategy for smart AI public policy. The EU could be very powerful in setting strong standards for auditing and measuring these AI systems in the wild, as currently there are very few standards at all. Creating these standards will require the perspectives of diverse disciplines and coalitions. And the process by which these standards are developed needs to be publicly accountable, academically rigorous and obviously subject to periodic review and revision.

Third, I think we need specific and actionable strategies to make the AI field more diverse and inclusive. I’d like to applaud the Commission for this very diverse example that you have here on this panel, but I have to tell you this is not what it normally looks like. Technical systems are of course sculpted by the views of the people who design them and there are policy levers that we can use to strengthen diversity and inclusion across the AI field and technology sectors. And we can support the participation of women, under-represented minorities and other marginalised groups. And obviously many are now recognising that the current lack of diversity in AI is becoming a major issue.

Fourth, and finally, I think there’s a powerful role for the EU to develop enforceable ethical codes to steer the AI field, after obviously extensive consultation and research. I was delighted to see that the European Group on Ethics in Science and New Technologies recently published a statement on Artificial Intelligence, Robotics and Autonomous Systems and I strongly support this important work in the creation of strong ethical standards in the sector.

However, in the face of rapid, distributed and often proprietary AI development and implementation, such forms of soft governance ultimately will face real challenges. Among these are problems of coordination amongst different ethical codes as well as questions around enforcement mechanisms that would go beyond voluntary cooperation. These are some of the discussions that are currently underway. New ethical frameworks for Artificial Intelligence need to move beyond models of individual responsibility to develop real accountability structures for industries and governments as they design and deploy AI. Otherwise, I worry that we run the risk of voluntary high-level ethical principles that can sound impressive but end up being used to substitute for consumer safety, fairness and democratic scrutiny.
Ann Mettler: Thank you. Professor Crootof, please.

Rebecca Crootof: AI is not responsible for the sins of the world, of course, but it enables the scaling of these sins. And that’s part of the problem: it enables them to occur on a new level. New technology always alters what human conduct is possible. Law and other forms of regulation alters what human conduct is incentivised or discouraged.

So I’m going to talk a little bit about legal responses to technological disruptions. There are basically three main ones: ban the technology; wait and see; and proactively regulate. All of them have benefits and drawbacks in different situations where they’re going to be more or less effective.

I’m not going to spend a lot of time on banning Artificial Intelligence for obvious reasons. Aside from the fact that it’s inherently dual use, scalable, and developments are likely to be rapidly diffused, we couldn’t ban it if we wanted to. And as evidenced by all of our discussion on the benefits of it, we probably don’t want to.

At the other end of the legal spectrum is the ‘wait and see’ approach, which is often celebrated for promoting innovation and avoiding overbroad regulation. I want to point out there’s no such thing as no regulation. There is just earlier and later regulation, and oftentimes later regulation can be far more reactionary and stricter and can hamper the development of beneficial systems even more than anticipatory safety standards, for example.

Furthermore, ‘wait and see’ approaches risk difficult- or impossible-to-remedy damage. We’ve had some analogies to big data as oil. I’ll just point out the last industrial revolution resulted in widespread worker injuries and pollution, and we do not yet fully understand what might be the negative side effects of an unregulated AI revolution.

But some of the harms discussed earlier foreshadowed some potential concerns. Proactive regulation is the Goldilocks solution: it attempts to address specific problems while still allowing space for innovation and creation. And I would say many of Professor Crawford’s recommendations would fall into that category of proactive regulation, targeted interventions.

It’s worth noting that direct regulation can be ‘tech-neutral’ or ‘tech-specific’. Again, there are benefits and drawbacks to both. Tech-specific regulations attempt to address a particular problem caused by a particular technology. So, for example, ‘sentencing algorithms must be interpretable and explainable so that individuals can challenge their results’ would be a very specific rule.

On the other side of the spectrum, tech-neutral regulations aim to preserve an overarching goal or principle, like a general requirement that machine learning datasets not amplify bias. The more tech-neutral a rule is the more flexible it is, the more future-proof it is. But with flexibility comes interpretability. So the more tech-neutral a rule, the more power that later interpreter has – be it the judge or the enforcer – to say what that rule actually means. The more tech-specific the rule, the more power to say what it means stays with the law creator or rule maker.

I’d also note, law can have an indirect impact in many ways: by incentivising certain actions such as industry practice through grant making, through tax breaks, through supporting codes of conduct, through requiring transparency. Law doesn’t need to be extremely heavy handed, it can be used as a nudge.

Because of its strong data privacy laws, the EU occupies a very interesting and potentially really productive regulatory position with regards to AI on the global sphere. It tends to have less access to data than other governments because of its privacy protections. But I think its activities are generally more trusted, it has a sort of higher ethical moral ground. Given this, I think the EU is uniquely well positioned to take the lead in developing ethical AI governance structures and I applaud you for having this event and considering how best to do so.
Ann Mettler: Thank you so much. Professor Dignum is next.

Virginia Dignum: Thank you. Indeed, the role of public policy is to ensure that AI is used for good and ensure the inclusion of all of the citizens and the diversity of the voices that are part of the discussion.

Regulations are, like I said at the beginning, potentially the key to achieve better solutions and better approaches to AI. We really have to invest not only in regulations that curtail or prohibit some types of activities – that is definitely important – but we should also look at regulations that encourage the use of responsible ethical approaches in AI. If we were egg farmers, we would have the choice of producing battery eggs or free-range eggs. By giving the incentives to be ‘free-range egg producers’ of AI, I think that Europe is in a position in which it can really show that ensuring the responsible ethical use and responsible and ethical results of AI is the way to go.

This requires that the population and citizens are educated and understand the difference between the two types of ‘eggs’ in AI and therefore it really goes back to the issue of providing and looking at how we have to deal with education. AI systems are actually systems which can give good answers. It is up to us and up to the way we educate ourselves and future generations which questions we put into those systems. And I think that that’s exactly what the change in the education system is: to train students to use questions, to create good questions and not so much to know the answers.

Finally, indeed the issue is – going back to the free range eggs – that responsible and ethical AI can be seen as the ‘new green’ and creating incentives for producing the green type of AI is something which we really can do in Europe and that we have proven in the past that we are very successful and can really do that.

Finally, concerning the opacity and black box algorithms, there are many systems that we don’t fully understand, including the way our own human intelligence works. That doesn’t necessarily mean and definitely doesn’t mean that we cannot control the limits of those systems and that we cannot regulate their applications. Transparency needs to be fundamentally concerned about the business processes involved in the development of AI systems. The processes which are used to maintain and govern, collect and train systems are the way we deal with the data and if we really can regulate this context in which the algorithm works, I think we, in Europe, are in a very good position to really control the effects of those algorithms even if we are not really able to fully understand how those algorithms work.

Ann Mettler: Thank you. Professor Fung.

Pascale Fung: So I’m invited here as the sole non-Westerner, so I’m going to provide my perspective as a non-Westerner.

If I were to record the conversations in Europe and European countries about AI and to calculate the frequency of words used, I would say ‘regulation’/’regulatory’ would be ranked very, very high compared to the conversations and dialogues in markets like North America or China.

So to begin with, to define your public policy and strategy, you need to define your priorities. Is it to be a leading AI superpower by a certain year, like China has set its own priority? Or is it to provide leadership in terms of ethical AI standards and ethical AI practices? Is it to provide fairness and inclusiveness in terms of AI applications? Or is it to provide leadership in terms of AI development and R&D. I cannot define that for you.

Having said that, I would say that public policy and strategy follow whatever you define as your priorities. You can use tax incentives, talent schemes, funding schemes and infrastructure investments, as well as regulations of course, to influence R&D, the development of AI, and the deployment of AI – its applications.
And I want to speak of some of the issues I see from our discussions here and elsewhere. In this context I want to speak about regulation first, because this seems to be of overriding concern to the EU. I want to caution that regulations have to be verifiable, quantifiable, enforceable and certifiable. Otherwise, it would just be talk, empty talk, and they will not have any effect on AI deployment. Because at the end of the day, what’s being deployed is AI technology.

I also want to caution against the demand for explainable AI for citizens. Number one: it is not possible for us to explain algorithms to all citizens. Number two: it is not necessary. Let’s look at the example of the drug approval process. I have no idea about the chemistry of the drugs I’m taking. I trust my doctor and I trust the FDA [Food and Drug Administration], you know, the approval body. So what we need to have for AI technology and to certify its safety and security is to have such an approval body consisting of experts and policymakers and whatever experts you want to have. And then, for the citizen to trust that system, because otherwise it is not possible. So that’s all I want to say about regulation, not at the abstract level, but at the concrete level.

Now in terms of R&D deployment: What I do not see in Europe compared to North America and China and other countries, is a lot of private sector research labs. Why is this important? The leading AI research is coming from leading AI companies in North America and in China. Why? Because the core algorithm development cannot have a leading edge unless it’s combined with data. So this is why it’s the Facebooks of the world, the Googles and the Alibabas and the Tencent companies. They all have their AI research labs. I do not see this in Europe.

So does that mean that we need to have such labs in Europe? Is it even possible, culturally and socio-economically, for Europe to have such big companies with research labs? I do not know, but if not, then Europe perhaps should focus on strengthening its core R&D capability, its leadership position – or potential leadership position – in the theoretical background of AI research, in ethical, I would say quantifiable, modelling of ethical principles, in cultural impact.

I don’t mean that at the abstract level but in really having a different kind of leadership role. So as not to compete directly against other countries that have certain strengths that Europe will never have. The lack of a homogeneous linguistic market is a huge barrier to data. Both North America and China have the advantage of having a single language data market. Does that mean that Europe should have a bilingual system where English is the lingua franca of user data? Is that even possible? Is that culturally or politically desirable? I do not know. But if we do not have it here, then Europe needs to consider what kind of leadership position it wants and go from there.

Question 6: In a nutshell...

Mario Mariniello: Thank you. We now approach our final question, for which you will have one minute only. Then the question is: In a nutshell: your message to the European Commission – what should (or should not) be done about Artificial Intelligence. And we start with Professor Steels.

Luc Steels: I think that the people, AI practitioners, researchers, are very happy that AI is now being discussed seriously at this level. So we are very grateful already for that.

Now I would say that a very important priority is to amplify European AI activities. Across Europe, there is a lack of knowledge – among scientists, developers and companies – about what is going on and the importance and actual impact on current technology. Greater recognition of European contributions is needed as that will attract investments towards European AI developments. It will also help to attract students to the field, bright students. And it will increase the acceptance by EU citizens because they will realise that AI is not only something coming from the outside – from the US or China – but is really being done by people in our community for the benefit of our community.
I think there’s fantastic AI work going on in Europe, in Italy for example, or in Portugal, Romania, Estonia and other countries... But it just doesn’t break through enough on the international level, because it is not known and due to the dominance of English-speaking media. It will never be known unless there’s some sort of action at the EU level.

Mario Mariniello: Thank you. Professor Crawford.

Kate Crawford: The AI industry is currently highly concentrated in the US and in China. But I think rather than believing in this narrative of a geopolitical race or in seeing this as a reason for fear, I think this can prompt us to ask a much bigger question. That question is: ‘What kind of world do we want to live in? And how can our technical tools serve that vision rather than driving it?’

To make those decisions well, we need more research – particularly interdisciplinary research, to understand the social implications of this turn to Artificial Intelligence. Without that, we simply won’t know how to maximise the benefits and minimise the harms as these tools continue to move further into our everyday lives.

I would suggest that it’s a false choice that the EU needs to pick between being a leader in AI development or being ethical and fair. In actual fact, the European Union can learn from the successes and mistakes of the early waves of AI growth, and it now has the ability to create conditions for a technology sector that is safe, fair and innovative.

The EU has an opportunity here to be the leader in AI governance. You can draw on the latest methods to ensure that, as the industry grows, there is a parallel process that supports public trust, due process, ethics and high standards of accountability.

I’d also like to thank you for bringing together this fantastic set of experts today and for hosting this event. It’s a very important time to have this discussion. Thank you.

Mario Mariniello: Thank you very much. Professor Crootof.

Rebecca Crootof: So there’s a common misconception often phrased as ‘law cannot keep up with technology’. I think this is both wrong and dangerous. I think it’s wrong because most laws apply to most new technologies most of the time. Those are not the cases that make the headlines, but most new technologies are governed by existing legal structures.

It’s also dangerous to assume that law can’t keep up with technology or can’t affect the development of technology, because it ignores the potential power of law to influence how technology develops and it ignores how a law can be used as an intervention to address harms that are identified. So I would say existing law already governs much of how AI is currently integrated into our lives. Simultaneously, AI is making the headlines because it is enabling new kinds of problematic conduct that are not yet regulated. We are confronted on a daily basis with examples of how existing law doesn’t address specific problems. And I think this is a space where a new proactive regulation is needed to address these harms.

I also want to say thank you so much for the opportunity to be here and speak today and for convening this group.

Mario Mariniello: Thank you. Professor Dignum.

Virginia Dignum: My main message to the European Commission on this issue is: we should embrace our diversity. We have discussed several times in this hearing today the importance of diversity and inclusion in AI in Europe.
We are indeed a diverse continent. We have different languages, different cultures. As a Portuguese person living for
the last 30 years in the Netherlands I can confirm all the difference across cultures in our continent. And I think that
this should apply as well to the way we look at AI and the development of AI technology and systems.

Let’s encourage taking the route less travelled. Let’s encourage the creativity which has been always a very
richness of Europe. Let’s look at different ways. We don’t have to follow; there are many different possibilities and
we really need to create the atmosphere and the conditions for different approaches. Not everything will succeed,
probably, but we need really to test that, to try many different things. As we have learned to live in a continent with
all our differences in culture and in including all those differences.

So I think that’s the main message and again I join my colleagues on thanking you for convening this meeting.

**Mario Mariniello:** Thank you. Professor Fung.

**Pascale Fung:** The recommendations I would give are to strengthen your investment in R&D, basic research,
and industry-academia collaborative research; create a single and more homogeneous marketplace and
foster collaboration with non-European partners.

You also need to strike a balance between doing all these things and realise that Europe still has an overriding
influence, a cultural influence, on the rest of the world – believe it or not, you still have a lot of influence. So, that
cultural leadership, perhaps you can take advantage of that because AI, unlike other technologies such
as wireless communications, is a very human-centric technology and as many of the panelists said, it’s
an interdisciplinary area.

So perhaps Europe can look into how using public policy to amplify its cultural influence in setting ethical standards
and good practice guidelines and so on and so forth, in working with the rest of the world in that area: How to use
cultural influence to spread the ethical use of AI?

**Mario Mariniello:** Thank you. Professor Luckin.

**Rose Luckin:** I would say embrace AI. It can bring enormous benefits, but as with everything with huge
benefits, it brings huge responsibilities because we need to ensure that those benefits are felt by the
whole of society, not just by a subset. And therefore, I think the Commission has a pivotal role to play both in
terms of leadership and action. And thanks for inviting me.

So, leadership through your AI-specific strategy that should be human-centered, should be framed
around empowerment, education and citizen control. Governance through an AI council to oversee all policy
initiatives, but sector-specific regulations to respect the context specificity of Artificial Intelligence applications.
And then action in terms of incentivisation from your procurement policies to capacity building, to research to
incentivising the development of good applications of AI, to the educational challenges that we all acknowledge that
we face.

**Ann Mettler:** We are now coming to the end and it’s now finally our turn to thank you. It was an incredibly
interesting hearing. I think you gave us much to think about and we are very, very, deeply grateful to you also that
you braved the taxi strike, which is very formidable.

So we now have come to the end of this hearing. We are serving an informal coffee outside of the room and I know
that there are probably many colleagues in the audience who want to have a word with you. This is probably a good
opportunity to do so. But before I let everyone go, maybe we can give a warm round of applause to our experts.
[applause]