Putting the OECD AI Principles into practice: progress and future perspectives

Session summary of this high-level panel discussion from the OECD MCM 2021

With the generous support
of the Government of the Republic of Korea

OECD 60 YEARS

Made possible by the generous support of

The Government of the Republic of Korea

TWO YEARS LATER, HOW FAR HAVE WE GOTTEN IN PUTTING THE OECD AI PRINCIPLES INTO PRACTICE?

A conversation to focus future efforts around trustworthy AI

Governments and other stakeholders have been working to implement the OECD AI Principles to make artificial intelligence trustworthy for the wellbeing of communities and individuals everywhere. This MCM side event was a timely occasion to highlight accomplishments and discuss future priorities.

The discussion showcased recent initiatives developed through the OECD.AI Network of Experts, including a framework for AI classification systems, recent reports on tools for trustworthy AI, national AI policies, and venture capital investment in AI. It was an opportunity to explain some of OECD.AI's most useful features such as live data about the supply and demand of skills on national job markets.

Participants discussed cooperative accomplishments such as the adoption of the G20 AI Principles in June 2019 that draw from the OECD AI Principles, the launch of the OECD.AI Policy Observatory in February 2020 and the work that its Network of Experts has accomplished. It also looked to the Global Partnership on Artificial Intelligence (GPAI) in June 2020 and the European Union's proposed AI regulation, announced in April 2021, and other international initiatives.

On a national level, the discussion highlighted how countries are implementing policies for trustworthy, human-centric AI, with in-depth presentations on how Korea and the United States have moved forward. It is worth noting that as of July 2021, 46 governments (38 OECD Members and 8 non-Members) have adhered to the AI Principles.

INTRODUCTIONS BY KOREA'S VICE MINISTER CHO

"As some of you know, the MCM is the OECD's key annual event for ministers from like-minded countries and partners to agree on our shared challenges, priorities and directions. This year's MCM marks 2 years since OECD members adopted the OECD AI Principles at the 2019 MCM.
As of today, 46 countries have adhered to these Principles, the first intergovernmental standard on AI, which subsequently served as the basis for the G20 AI Principles.

The 2019 OECD AI Principles represent a shared vision/ that helps us steer AI to benefit people and planet by promoting the use of AI that is innovative, trustworthy, respects human rights, the rule of law, and reflects open and transparent market-economy principles. I am proud to say that Korea played an important role in the development of the AI Principles, including through chairing the Committee on Digital Economy Policy and the expert group that developed the Principles.

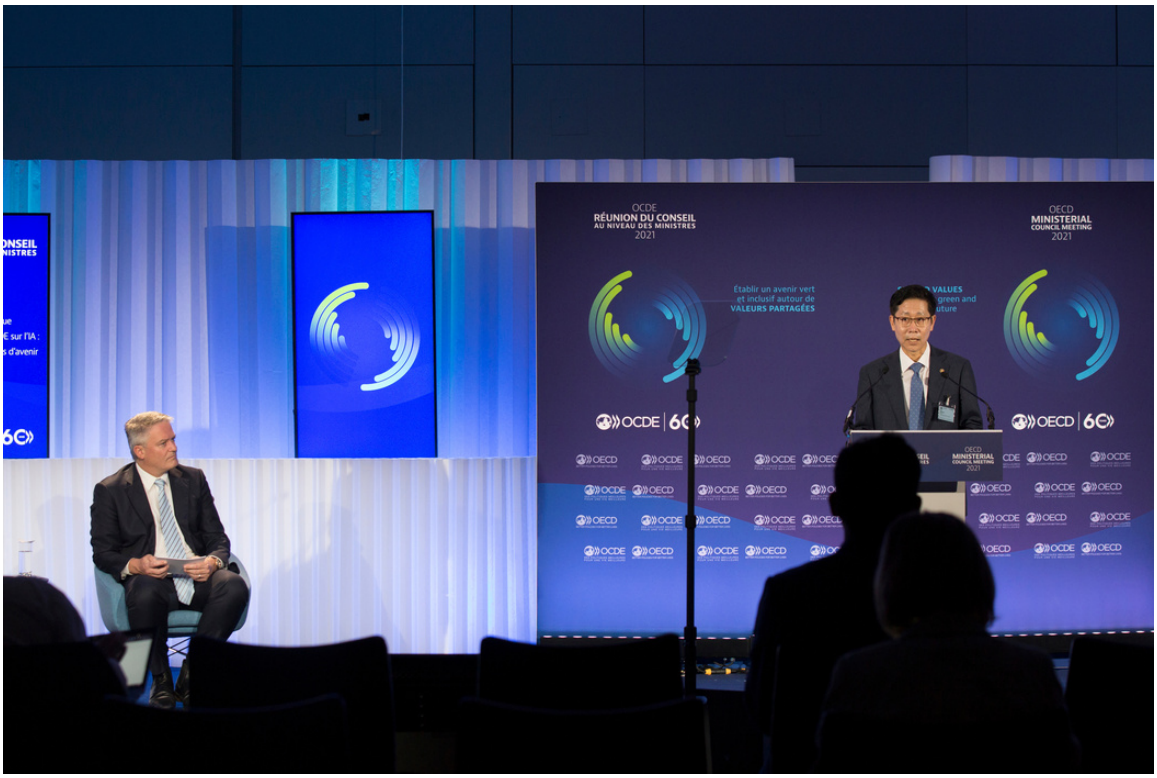
In terms of implementation, three key aspects of the OECD AI Principles include "developing a measurement framework for AI"; "monitoring the implementation of the Principles" And "providing a multi-stakeholder forum to discuss AI policies."

Since their adoption, the OECD has surpassed our expectations and broken new ground in order to help all OECD member countries and many others to share and shape their AI policies through the OECD.ai platform that provides: live data on AI development; a database of national AI policies in over 60 countries; fruitful collaboration with AI experts globally, as well as on collaboration with other intergovernmental organisations working on AI.

I can think of no technology more critical to the future of our economies, of our societies, and of our environment than artificial intelligence. That's why Korea organised today's event with the OECD.

Two years after the adoption of the OECD AI Principles, this discussion will highlight how our countries are implementing policies for trustworthy, human-centric AI.
The 3 things I would like to suggest we focus on today are: First, Lessons learned to date in implementing AI policies to steer the transformation of our economies and societies for inclusive growth and addressing global challenges. Second, AI governance and soft and hard law, including the proposed EU AI Act which are triggering discussions around the world, seeking how to ensure that our policy and legal frameworks for AI are interoperable globally so we can maximise the benefits and minimise the risks of this powerful technology.
And third, what are the priorities for international cooperation on AI moving forward and how can the OECD help us with this?

I look forward to the presentations and panel interventions to discuss and share different perspectives on these three priorities."



INTRODUCTIONS BY OECD SECRETARY-GENERAL MATHIAS CORMANN

“Thank you vice-Minister CHO for initiating this timely event and for your kind words about the OECD’s work helping us shape and share policies for Trustworthy AI that benefits people and planet and monitor our collective progress.

My OECD colleagues will detail our work on the policy and regulatory front, but I’d like to use my time today to talk about the AI ecosystem today and share some findings from three new reports that are fresh off the printing press.

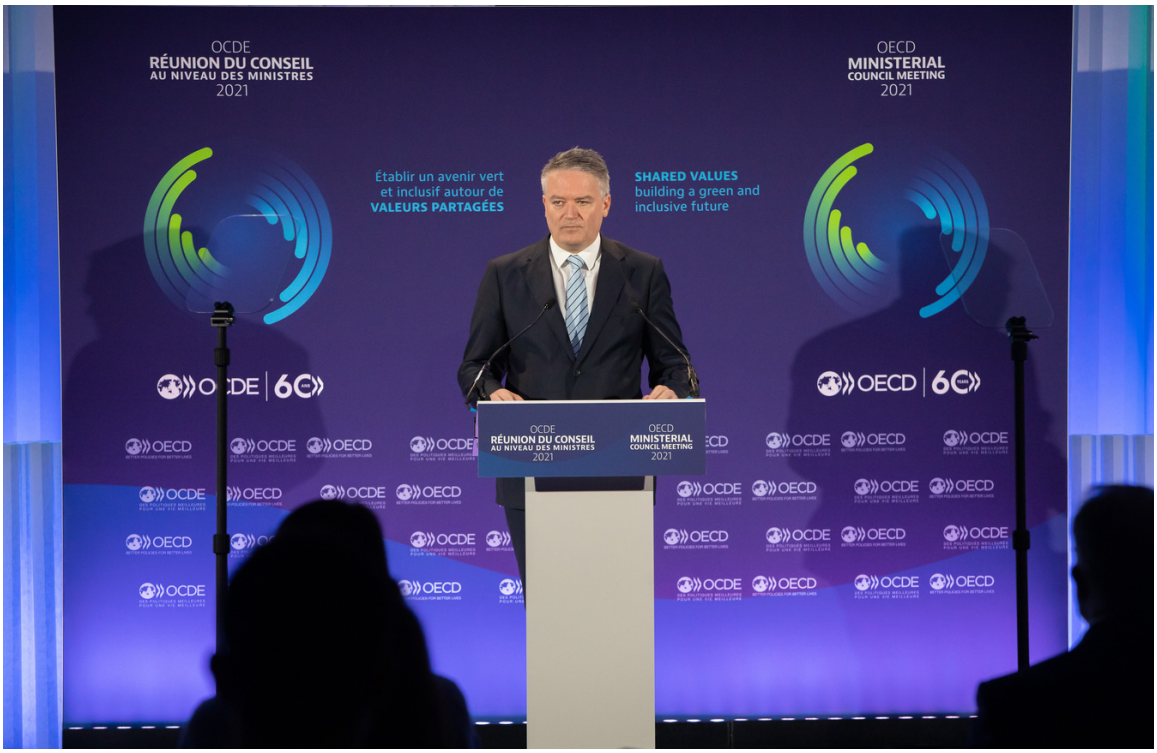
OECD ‘dot’ AI helps us track the most recent developments and trends in the AI ecosystem in areas such as AI policy, software development, research, investments and the demand and supply of AI skills. Global AI research over the last five years has grown by 36%[i]. Demand for AI skills over the last three years has also grown; by 50% in countries like Ireland, Spain and France; doubling in the United States, Australia and the United Kingdom; and almost tripling in countries like Canada, India, Indonesia, and Brazil. Today, a job opening is one and a half times more likely to require AI skills like machine learning than 3 years ago[1]. And in just 5 years[ii] AI software development[2] around the world grew by a dramatic 900%!

The OECD “dot” AI Policy Observatory is constantly adding new analysis and data. And today is no exception: Our new report on AI venture capital investments in AI from 2012 to 2020 is full of findings about who is investing in AI-driven private companies and where these investments are going. In 2020 alone, VC investments in AI represented more than USD 75 billion – and this is just the private companies! We’re also working on analysing public companies’ investments. AI-driven companies now account for over one-fifth of all VC investments. The United States and China lead the way – together attracting 80% of venture capital in 2020, followed by the EU, Britain and Japan, with the US over half of this amount[iii]. In terms of sectors: transportation leads investments, pointing to our driverless future[iv]. Second are investments in health, pharmaceuticals and biotechnology that skyrocketed in 2020 with the pandemic and with increasing demand for personalised healthcare. And third are business process tools that reflect the role of AI in increasing efficiency and productivity in the workplace.

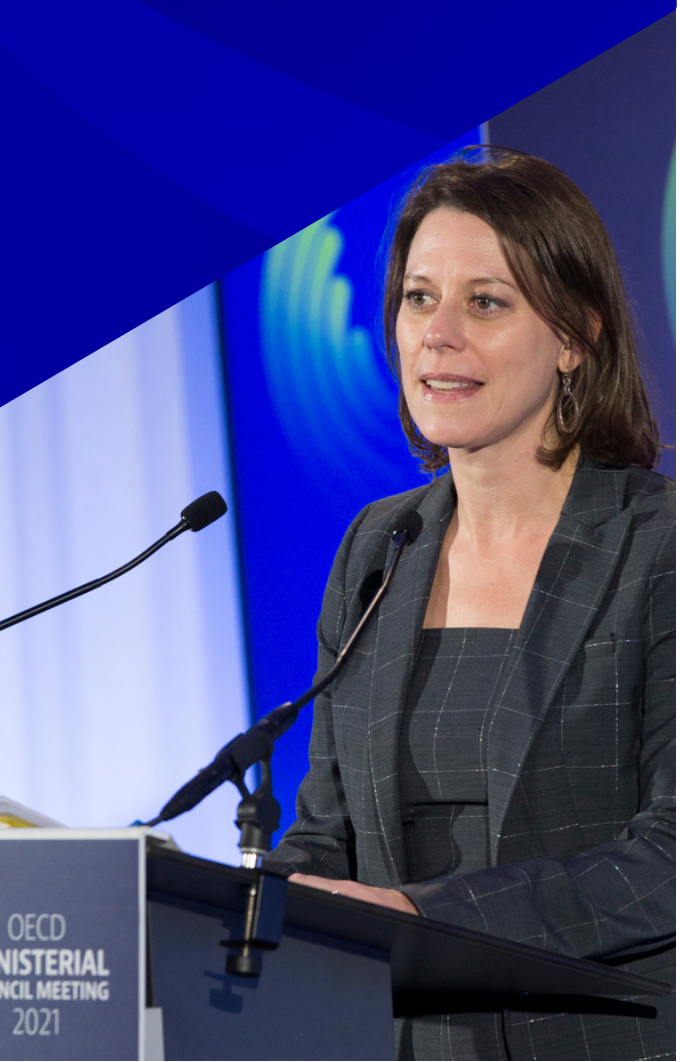
This leads me to our work on AI and work, innovation, productivity and skills, where we looked at who develops AI-related innovations, goods and services through patents and trademarks and see the same trends, with significant growth since 2014 of AI-related patents and trademarks. The 5 companies with the largest AI intellectual property portfolios own 14% of the total IP portfolio related to AI, pointing to market concentration in the field.
We looked at the current and future AI jobs and at the skills we need people to have so they are able to leverage AI. Machine learning and data skills are critical for new jobs today. So too are teamwork and socio-emotional skills, and being able to ask the right questions so AI can solve the right problems – these skills are inherently human.

While the OECD develops evidence-based public policy recommendations in a multi-stakeholder fashion, cutting-edge research and pilot projects on AI priorities are also critical to advancing trustworthy AI. For this reason, we are hosting the Secretariat of the Global Partnership on AI, which is an international initiative originating in the G7 that promotes responsible AI use through applied co-operation projects, pilots and experiments.
Finally, the OECD's reach on AI policy extends well beyond OECD countries, Including through our close cooperation with other inter-governmental organisations with complementary mandates and shared values on trustworthy AI, human rights and democracy.2 weeks ago we launched Globalpolicy.AI – a new online platform that brings together eight partner intergovernmental and regional organisations. This online platform aggregates our organisations' work on international AI policy to create a one-stop-shop for AI policy experts and the wider public.

I wish you a very productive session today."



Progress over the past two years in implementing the OECD AI Principles



OECD.AI: RESOURCES FOR TRUSTWORTHY AI

To kick off the first session, OECD Deputy Secretary-General Ulrik Knudsen introduced a video that presented OECD.AI and its main features.

After the video, Audrey Plonk, Head of the OECD's Digital Economic Policy Division presented the OECD's work over the last two years focusing on artificial intelligence policy.

National AI policies, best practices and future indicators

She explained that the AI policy cycle depends on a working group of national AI policy coordinators who leverage OECD.AI and its database of AI policies. The database covers AI policies and initiatives from more than sixty countries. The working group looks to the database to identify good practices for implementing the AI Principles in the areas of knowledge and skills, R&D investment, regulation and documentation, but also tools for trustworthy AI.

The working group also identifies policy intelligence in action plans and targets, while evaluating how countries implement their policies through benchmarking. The group is also working to establish indicators for trustworthy AI. Ms Plonk pointed attendees to a recent report at oecd.ai/policies that compiles these findings. She reiterated her team's commitment to collecting more data about national policies and implementation in the future.

The OECD Classification Framework for AI Systems

Ms Plonk presented a second aspect of the Secretariat's work, the AI system Classification Framework for assessing AI systems and applications. She explained that the framework is meant to help policy makers assess the risks and challenges that the wide array of AI applications and systems present.

This user-friendly framework will be launched early in 2022 and looks into four dimensions of AI systems: the system's context; the data and input that goes into the system; the AI model and its building process; the task and output and its level or degree of autonomy. The tool will be available at oecd.ai/classification and it will be applicable to AI systems across the board.

To date, the Framework has been applied to hundreds of AI systems and applications. She explained that it will also be the basis of a risk assessment tool for AI being developed in cooperation with partner organisations such as the European Union.

Tools for trustworthy AI

Finally, Ms Plonk presented the Secretariat's work on identifying tools for trustworthy AI. A recent report available at oecd.ai/tools maps the tools that the working party has identified to three main groups. The first is technical tools such as standards, software and research. The second is procedural tools for things like risk management and governance framework tools. The third regroups educational tools for stakeholders involved in or affected by AI systems.

She announced that there will soon be a catalogue of trustworthy AI systems available on OECD.AI. The catalogue will categorize the tools based on their characteristics in a friendly searchable way designed for policy makers.

Looking forward, she said that her team will explore mechanisms to ensure that accountability for trustworthy AI, such as technical or procedural standards, impact and conformity assessments and certificates, are widely established.

The OECD.AI Network of Experts

Ms Plonk explained that all of the OECD's work is informed by experts and partner organisations to focus on different aspects of trustworthy AI. These include the OECD.AI Network of experts that provides policy guidance, the Global Partnership on Artificial Intelligence, and the OECD Parliamentary Group on Artificial Intelligence.

New data sets around the AI ecosystem

Ms Plonk announced new data sets that will be released in the coming months, including live data on venture capital AI investments mentioned earlier by the Secretary-General. The data sets will be linked to help understand the different ingredients of the AI ecosystem. For example, one could look at AI research and skills, data input and investment capital but also computing capabilities. She explained that this information will help to understand which are the best recipes for policies in each country.

AI legislation proposals and an incidents database

The OECD is also expanding the national policy data on OECD.AI to track legislative proposals on AI. Many countries are advancing legislative proposals, so it is interesting to track and understand the proposals to see what works well.

Finally, she announced that the OECD will develop a database of AI incidents, hosted on OECD.AI. This incidents portal will pool information about AI incidents from all over the world to create consistency and comparability in reporting, measuring and analysing such incidents. This is critical for policymaking and feeds back into the overarching concept of risk assessment that is present in many of the OECD's AI activities.

THE EUROPEAN UNION'S COORDINATED AI PLAN AND PROPOSED LEGISLATION

Moving on to the next presentation, Deputy Secretary-General Ulrik introduced the next speaker, Mr Roberto Viola, who is the Director-General of the Directorate General of Communication, Networks, Content and Technology (DG CONNECT), of the European Commission.

This April, DC CONNECT released an AI package that includes the EC review of the coordinated plan and also a proposal for the AI Act.

Mr Viola then gave a brief overview of the main components of the package and how it aims to promote the development and deployment of trustworthy AI systems in the European Union and next steps.

First Mr Viola took a moment to say that he feels that the European Commission's work on AI is very close to the OECD's thanks to close cooperation. The OECD's AI Principles played a key role in forging the EC's AI package that was presented in April.

He explained that the EC's package is based on two pillars. First is the ecosystem of excellence, which is what Europe needs to foster research and investment in AI. Then there is the ecosystem of trust. He stated that the European Commission wants all stakeholders to trust AI because they see it as a gift for humanity and it can do many wonderful things, particularly when it comes to science, new drugs and understanding the pandemic.

A risk-based approach to regulation

However, he went on to say, that government stakeholders should start building upon the work of the OECD and start creating some rules into regulatory systems. The EC's initial approach is that not everything needs to be regulated, in fact, only a minority of AI applications should need regulation. The Commission's approach is based on risk, where one looks at risk based on a pyramid, where the pinnacle represents things that should never be allowed, such as social scoring by public organisations, or manipulating minors and people with mental disabilities.

Then there are a number of scenarios where AI is operating in risk-filled environments where human life is at stake, such as a car or flying a plane. Also, there are systems in which our right to have a future can be placed in danger. For example, in systems that grant loans, there can be unwanted bias by design, whereas it is a fundamental right in democratic societies for everyone to have a fair chance. There are many examples where this could be at risk and we want to be sure that is not an issue. That is why AI should favour access to work, a loan, schools and to healthcare while managing any risks. Mr Viola reminded attendees that the European Commission's risk-based approach is horizontal in nature as opposed to sectorial.

Risk assessment should plug into all legislation, and the EU hopes that this horizontal, risked-based approach can serve as a model for others outside of the EU. Mr Viola stated that the EU would be happy to work with other like-minded governments to help them to integrate this approach.

International interoperability of the EU's approach to AI

Mr Viola stated that interoperability is a must because no one can advance in AI in isolation because our world is interconnected and it is a common interest to ensure that what we do works together. He went on to say that the EU is very open to deepening its approach to risk-based regulation and cooperation, as it already has done with the OECD on a number of projects, such as supporting the Secretary-General's announcement about work on risk assessment.

Mr Viola pointed out that the EU is also open to working bilaterally with other countries, and thanked the United States for the Trade and Technology Pact that was inaugurated in Pittsburgh a week earlier. With the United States, the European Union agreed that they should work on risk assessment and the implications of regulations, but also standard and norm-setting for machine learning. The work happening on the trade and technology panel explicitly mentions the OECD as the multilateral way forward. He acknowledged that this was one example but that the European Union is working with many friends around the world on AI-related issues, including work with the GPAI.

THE UNITED STATES' NATIONAL AI INITIATIVE

Deputy Secretary-General Ulrik thanked Mr Viola and introduced Dr Lynn Parker, Director of the National AI Initiative Office at the Whitehouse and Assistant Director for AI at the Whitehouse Office of Science and Technology Policy. Deputy Secretary-General Ulrik set the scene by stating that the National AI Initiative Act of 2020 became law early in 2021 and asked Dr Parker to tell listeners about its main components.

Dr Parker took the floor to explain that the objectives of the National AI Initiative Act are fourfold. First, the initiative aims to ensure continued leadership by the United States in R&D and development. Second, the United States would like to lead the world in the development of trustworthy AI systems in the public and private sectors. Third, the United States must prepare current and future workforces for the integration of AI systems across all sectors of the economy and society. Finally, the United States must coordinate AI R&D and demonstration activities among federal agencies to ensure that each informs the work of the other.

She continued by explaining that the legislation also created the National AI Initiative Office in the Whitehouse, which acts as a coordinating hub across agencies to drive towards the initiative's goals. She said that the Office is pursuing policy initiatives in a number of priority areas. To elaborate, she went on to describe four of those initiatives: R&D – including R&D infrastructure, trustworthy AI and education and workforce development.

AI R&D including infrastructure

Prioritizing AI R&D is happening across a number of specific R&D efforts. One major example is the Office's goal of building up larger scale and multidisciplinary research in AI. The United States has launched eighteen large scale, multi-partner research institutes across the nation. They are led by the National Science Foundation along with other agencies and private sector partners. So far, investments in these institutes total over 360 million dollars over five years. It is important to note that these new AI research institutes are partially funded by the private sector, which she said is key to creating an AI ecosystem that is dynamic, innovative and in line with American values.

The United States' international priorities and the OECD

Deputy Secretary-General Ulrik then asked Dr Parker to discuss the United States' international priorities and how the OECD fits into them. Dr Parker responded by saying that international cooperation in AI is also a strong priority for the United States and that it deeply values collaboration with like-minded allies to advance the development and use of trustworthy AI. She emphasized that acceptance of the OECD AI Principles in 2019 was a watershed moment for the United States that all can be proud of, and that is a sign of international commitment to foster trustworthy AI. She outlined the OECD's important role in this effort, especially when it comes to bringing together multiple stakeholders with a breadth of expertise, diverse experiences and backgrounds to advance a shared understanding of policy approaches for AI. The United States has been a strong supporter of the OECD.AI Network of Experts on AI and supports the OECD's Digital Economic Policy Committee's decision to create a working party on AI and suggested that this would solidify the OECD's position as the premier forum among democratic nations to discuss AI policy. She continued by expressing the United States' enthusiasm for the ongoing work on AI computing capacity and named the OECD as the United States' preferred venue for discussing AI policy on an international scale while supporting the GPAI and its close collaboration with the OECD to bridge the gap between theory and practice in AI policy.

She concluded by saying that the United States is committed to working with all of its international partners to ensure that AI lifts people up, solves problems and advances human freedoms, and thanked the Republic of Korea and the OECD for hosting the event.

Next, Dr Parker explained that strengthening the national AI infrastructure is also a priority. A key initiative in this area is a concept the National AI Initiative Office calls the National AI Research Resource. Its strategic objective is to strengthen the United States' AI innovation ecosystem by supporting foundational research in AI, advancing youth inspired AI research to accelerate discovery across many fields and to increase the number. It also aims to increase the number and diversity of researchers participating in AI research. This resource is envisioned as a shared research infrastructure that provides researchers and students across all scientific disciplines including computational resources, high-quality data educational tools and user support. A task force is working on a roadmap and implementation plan for this resource, which will be delivered by November of 2022.

Trustworthy AI

The third priority that Dr Parker discussed is to advance trustworthy AI. To illustrate, she presented a newer effort by the National Institute of Science and Technology (NIST) to create a risk management framework. NIST is engaging with many stakeholders after issuing a public request for information on the framework and the comments it received are available online and held a public workshop in mid-October. Dr Parker encouraged anyone listening to engage with NIST on the framework so that they could obtain a high-quality result that is helpful to the entire AI community.

An AI-ready workforce

Finally, Dr Parker went over the United States' priority to train an AI-ready workforce, to fill the AI talent gap and preparing United States' workers for jobs of the future by implementing policies that ensure a diverse, inclusive and knowledgeable workforce. She continued by stating that the United States would like to see AI-related concepts integrated into all levels of schooling from prekindergarten to post-doctoral positions and including community colleges. Coalitions of private sector, non-profit and government partners can help to obtain education objectives to create effective AI education at scale.

Dr Parker encouraged those who would like to know more to visit ai.gov or to go to the OECD.AI Policy Observatory for the latest updates.

KOREA PREPARES FOR THE AI TRANSITION OF ITS ECONOMY AND SOCIETY

Next, Deputy Secretary-General Ulrik introduced Kyunghee Song, the Director-General of the Artificial Intelligence Policy Bureau of Korea's Ministry of Science and ICT (MSIT) and asked her to inform attendees about how Korea is preparing for the AI transition of the economy and society, and what Korea's priorities for international cooperation.

Ms Song began by introducing Korea's progress in implementing the OECD AI Principles over the past two years and indicating its future directions. She explained that AI has emerged as a key technology and innovation driver that has fundamentally changed industries and lives. To steer AI to become a key driver of the fourth industrial revolution, Korea developed a strategy for artificial intelligence in December 2019. Then, to overcome COVID-19, recession and accelerate the digital transformation, in July 2020 Korea launched the Digital New Deal including the data and products. Korea's digital strategies are closely related to the OECD AI Principles and their five policy recommendations.

A strong digital ecosystem for AI

She explained that Korea first committed to strengthening its digital ecosystem through data and products. This effort aims to stimulate all stages of the data lifecycle from data accumulation to data use to address the shortage of data in the public sector. She said that Korea will build a data dam by acquiring high-quality data for AI training and making public data accessible.

Korea will also facilitate use of the acquired data by establishing an AI hub to provide companies and researchers with AI training data from the dam and cloud-based high-performance computing. The ecosystem will include big data platforms to produce and manage data, especially for sectors such as finance and healthcare. Korea will also provide the companies and researchers with the data and AI vouchers for supporting usage fees.

n addition, Korea will develop an AI health product aimed at promoting AI adoption by using the large volume of data housed in the public sector to support the development of AI services. For example, the Korean government is providing companies with experimental labs equipped with the data, computing power and memory to support medical imagery analysis to improve the quality of medical services in the military or to facilitate the COVID-19 epidemiological investigation.

A policy environment that enables AI

Ms Song then went on to explain that Korea is fostering an enabling policy environment for AI. Just last year, Korea amended its three main privacy laws to promote data use and enacted a framework act on intelligent informatization to foster an enabling environment for AI use. Korea also established a roadmap for revamping laws, systems and regulations and access guidelines for AI. Ms Song explained that even more works are underway to overhaul law systems to first enable the environment for human-centred AI.

Education and training to prepare the workforce for AI

For her third point, Ms Song said that Korea is carrying out educational and training programmes with the aim of building human capacity and preparing for labour market transformations. AI talent is being nurtured to build human capacity through measures such as increasing the number of AI graduate schools and operating short-term intensive educational programmes like the Innovation Academy. In addition, basic education programmes for improving AI skills have been provided to all citizens by operating software schools and AI-oriented schools and making online education content available. This will help citizens boost their capabilities to adapt to emerging technologies in response to labour market transformation.

Research and development in AI

For her fourth point, Ms. Song discussed Korea's efforts in AI research and development. Korea is investing to develop AI semiconductors such as processing in memory (PIM) and a Neural Processing Unit (NPU) in cooperation with the private sector and universities.

Korea is also working to secure next-generation technologies such as AI which has an explainable decision-making process and AI that can be trained with a small amount of data and AI with applied learning.

Korea among the most AI-ready countries

Finally, she reminded attendees that Korea has been proactive in international cooperation for developing trustworthy AI. In May 2019, as Vice-Chair of the OECD Meeting of Council at Ministerial Level (MCM), Korea contributed to the development of the OECD AI Principles to impose social responsibility for AI. Korea is currently involved in developing UNESCO's Recommendations on the Ethics of AI. Korea also hosted the Asian Pacific Regional Consultations in 2020 to help get stakeholder feedback in the region.

Ms Song finished by saying that thanks to such proactive policy implementations and the government's AI readiness in 2020, lauded by Oxford Insight, Korea ranked seventh out of 120 countries, which is up nineteen places from the previous year. The implementation of Korea's national strategy for artificial intelligence and the Digital New Deal seem to have had a positive impact on Korea's upward climb in the ranking.

Ms Song acknowledged that this event had been a good opportunity for Korea to show the progress it has made in implementing the OECD AI Principles and to agree on the necessity of international cooperation. Ms Song then invited all to participate to follow Korea's as it moves towards the goal of advancing AI technologies and industries for the benefit of all human beings

Korea sees the OECD as an excellent platform to share best practices on AI policies

Next, Deputy Secretary-General Ulrik thanked Ms Song and asked her where she thinks the OECD can help most to help make progress on this agenda.

Ms Song replied by saying that she thinks the OECD is providing a great platform where states can share cases and best practices of AI policies that can be used for benchmarking to deal with the global challenges that governments face together. She went on to praise the OECD by saying that she is always impressed by its various activities including the recently established OECD.AI platform but also the recommendations and principles that it has set out.

She said the most important thing in international cooperation in AI today is how and what kind of regulatory framework can be discussed in AI development. She went on to add that regulations that do not take into account that technological developments can cause too heavy of a compliance burden for SMEs, and that can ultimately hinder innovation.

She continued by saying that market fragmentations may also happen depending on the intensity of regulations which can be obstacles for free trade and create a detrimental loss of benefits for all parties involved. She said that there is a need to focus on establishing trust in AI and discuss both the good and the bad on an international level, and it is important for all of the interested parties to be involved in establishing trust. She closed by stating that the amount and the kinds of regulation required to ensure trust in AI still needs to be discussed and expressed appreciation for the OECD's efforts in this area.

QUESTIONS FROM ATTENDEES

"We've heard a lot about domestic R&D efforts. Could you tell us about international plans in research and development, notably for trustworthy AI?"

- Marko Grobelnik, AI Researcher at Jozef Stefan Institute

Dr Lynn Parker answered by saying that there are many international agencies and collaborative efforts that are involved in research and development. They sometimes take the form of joint programmes from one nation to another and address many of the critical challenges that still need to be solved. There are also personnel exchanges happening to allow each side to learn about the other's programmes and approaches. She also pointed to joint workshops and similar activities. She then mentioned the robust collaboration happening between the United States and the United Kingdom. She finished by saying that the United States is engaged with countries all across the board and is willing to grow.

Mr Viola also answered by saying that the EU welcomes all forms of international cooperation and underlined the fact that it is the basis for the EU's R&D programme. He went on to say that the EU now has what is called the Digital Europe programme for testing AI. This programme reaches out to countries around the world to cooperate and associate with that programme. The EU is also working bilaterally with some countries and has begun some pilot projects and is very much open to the idea that the more the EU works with like-minded partners, the more everyone will advance on trustworthy AI.

"The first AI principle talks about inclusive growth and sustainable development. How do countries like Korea plan to adhere to this promise and include low income and low skilled workers in the prosperity that AI is expected to bring?"

- Nozha Boujema, AI Director at IKEA

Ms. Song responded by acknowledging that this is a very challenging question and went on to say that Korea sees AI as the technology leading the fourth industrial revolution but they are focusing on the transformation in labour markets and the fast changes.

To respond to this transformation, Korea has been educating low skilled workers by providing various opportunities for training and solutions that are fit to different levels. The Ministry of Science and ICT offers these training solutions along with the Ministry of Labour. They also look at the population as a whole to promote advanced AI products, because there is no point in developing technologies if no one can use them.

Ms. Song acknowledged that the gap between the digital haves and have nots is widening but said that Korea is trying to ensure that the gap does not grow bigger, because Korea believes that AI can help everyone, regardless of social status. Korea also looks to international activities to find better responses to this challenge.

TRUSTWORTHY AI FROM AN INDUSTRY PERSPECTIVE

Mr Pilat introduced Mr Eric Horvitz, Chief Scientific Officer at Microsoft, where he is responsible for trustworthy AI, co-founded the Partnership for AI, and created Stanford's One Hundred Year AI Study that focuses on AI in the long term. Mr Pilat asked him to explain what he sees as trustworthy AI from an industry perspective.

Mr Horvitz took the floor to explain that a key part of competition when it comes to a new set of technologies such as AI is to understand that there are unexplored influences on people and society and that earning trust can be quite competitive. He went on to say that in his experience, industry often finds itself at the translation where the core AI advances into the complexity of the open world, realistic applications. So industry R&D often gets early exposure to challenges in trustworthiness including concerns about robustness, safety and fairness, and of course the cost and benefits of using AI out in the open world. This happens in advance of regulation and government interest.

He went on to say that competitive pressures notwithstanding, leading technology companies are responsible for taking a serious look at the issues for their customers and the greater public. His top priorities for developing trustworthy AI in industry are:

1/ invest in understanding the benefits and cost of the technologies;

2/ work to mitigate concerns and make decisions about if and how to ship technology;

3/ develop transparent documentation about AI systems and services that capture limitations and also promote best practices;

4/ develop tools and methods for understanding AI systems and methods across the whole AI lifecycle, including the maintenance of systems over time;

5/ engage and dialogue with civil society and government on their findings;

6/ help customers to pull everything together and establish a set of guiding principles and cross-company programmes.

THE HIGH RISK AI APPLICATIONS AND CLASSIFICATION SYSTEMS

Next, Mr Pilat introduced Andrea Renda, Senior Research Fellow and Head of Global Governance, Regulation, Innovation and the Digital Economy (GRID) at CEPS.

He asked Mr Renda to share his perspective on the issue of high risk AI applications and classification systems and any approaches he thinks are adequate for dealing with the risks they present.

Where AI risks occur

Mr Renda said that most experts acknowledge that any approach to AI policy should be risk-based. Implicitly, this means that different ways to design and deploy for different use cases create different risks. Defining high-risk AI is, therefore, a very important but also challenging question. Generally, he said, there are two answers to that question. One is easy and the other is more complicated and difficult to digest.

We are broadly converging on the idea in the policy community is that high risk AI acts in two main areas. One is the domain of fundamental rights, human rights, not only privacy and discrimination but many other personal liberties and rights including freedom of expression and human agency, start businesses, freely assemble and many others. Looking at what has already emerged from existing AI systems, this area is generally agreed to be high risk whenever there is a chance of infringement upon these rights. The other area is that of safety and security, particularly when it comes to managing critical infrastructures. This is both because it is AI and because it is software-based. The authorization of specific functions often comes with specific risks. This brings on the possibility that deploying an AI system within a broader more complicated system generates specific risks.

He continued by reminding attendees that these are the easy answers emerging from the EU AI Act and to some extent the OECD AI Principles, but it is clear that risk is a context-specific issue. The way AI is designed but also the decision to deploy it, the governance arrangements and the risk mitigation arrangements that accompany it might determine the extent to

which a system might be considered high risk or not. He noted that this is a much more challenging and interesting area where work is just beginning in the OECD and focusing on use cases but it is the beginning of a long journey. He said that everyone needs to learn about what the risks are and the potential government arrangements that produce the potential risks in an AI situation and what the key pressure points and decision-making moments are that need to be examined.

Design, deployment and human oversight

One last thing that Mr Renda considered to be important in the simple response, is that actors have come to realize that it is not sufficient to look at factors when the AI system is designed. It is equally important to look at the position of those who deploy the system and decide which data to use and which human oversight arrangements to adopt and all of the governance and organizational features that will accompany the system when it is placed on the market and after the AI system has reached the market.

Next Mr Renda turned to the more complicated yet brief response. He said that for the time being no one knows yet and probably will never fully know exactly what high-risk AI systems are because it is a very fast-changing technology. This means that policy makers must understand that the key moment in policymaking on AI is the governance set up to monitor AI in the market and to monitor interactive effects, particularly when it interacts with the external environment and other AI systems. Here, he said, is an area where collaboration should be stepped up between policy makers at the OECD and elsewhere.

International cooperation for standards and R&D

Mr Pilat then asked Mr Renda to explain where he sees that the priorities lie when it comes to international cooperation on AI.

International cooperation is essential to share resources and experiences. It can already be mapped out in a number of fields, certainly in the area of regulation, definitions, risk classification systems, risk assessment models, and standards will all be important areas going forward. Standards-setting is a very global effort for the moment and it must remain as inclusive as possible.

R&D projects are also an area where international cooperation is important. He also stated that AI has the potential to do wonderful things for humanity but this will only happen if policy makers give direction and join forces to deploy AI in support of the global public good.

To make this happen, he said, we need to be as inclusive as possible and echoed Ms Song's earlier statement that all actors must ensure that AI does not exacerbate already existing inequalities.

At the OECD, Mr Renda said that he would like to see a commitment between member countries, working in the regulatory committee, to consider international regulation on AI when legislating at the national level. This means they should avoid legislating obstacles to international cooperation, data sharing and R&D collaboration when adopting national initiatives. He remarked that this is already happening in Canada, the United States and the EU, but he stressed that it should be formalized. He closed by saying that only international cooperation can help to maximize the benefits and mitigate the risks that AI will bring in the coming decade.

THE NEED TO DEMOCRATIZE ACCESS TO AI AND DATA

Mr. Pilat introduced Mr. Vilas Dhar, President and Trustee, Patrick J. McGovern Foundation, where he advances data and AI solutions to create a thriving, sustainable and equitable future for all. He asked Mr. Vilas to talk about the need to democratize access to AI and data, and to ensure that AI developments benefit individuals and communities.

Mr. Dhar opened by going back to the human story and looking at what it means to create AI for the world. He asked attendees to think about what it means for policy makers to make frameworks for new actions, both in an industry and a government context. But he insisted that AI is really transforming the human experience. He went on to share three applications that the Patrick J. McGovern Foundation is currently engaged in.

A few of the McGovern foundation's projects

First, Mr Dhar pointed to the use of natural language processing, not just for the languages commonly used across the OECD and the UN but also to show how the languages of indigenous people and ancient or dying languages are represented in a global construct. This should help to show how AI can create more inclusive participation in political and commercial processes to make sure that everyone is actually a participant in the digital economy.

The second application that Mr Dhar discussed falls into the context of climate change. These are applications that allow individuals who are facing the impact of climate change in their communities to report what is happening to city, regional or state lawmakers not to be able to understand the massive data inputs to create policy responses that operate at regulating the technology. Rather, to build new infrastructure or new direct relief mechanisms that will allow the much needed transformations.

The third application works on a macro level. In reference to the 15 trillion+ dollars that might be added to the global economy by 2030 thanks to AI applications, the Patrick J. McGovern Foundation is focusing on subsistence level farmers and how agricultural technologies might be applied so that individuals can better understand planting and harvesting technologies, and to receive better pricing when they take their goods to the global economy.

He went on to say that these three examples are not just a catalogue of AI applications today. They help to consider how policy makers might think beyond the challenges presented by current applications of AI. It is more about incentivizing positive, human-centred approaches in the years ahead.

When it comes to R&D funding, he said, the Patrick J. McGovern Foundation wonders, as the funds flow both to academia and industry to build these kinds of applications, how is it possible to ensure that they are also flowing to civil society organisations? After all, civil society is in the best position to know how these applications can actually lead to transformative effects?

Three approaches to human-centred AI

As one of the leading voices in the civil sector, advocating for a new, human-centred approach for AI, the Patrick J. McGovern Foundation seeks three approaches. First, by giving direct funding and support for non-profit organisations that work at the intersection where policy meets commercial trends. Last year, the Foundation shared 40 million dollars and will potentially share 75 million next year, with non-profit organisations globally. The financial support goes to more than the development of new products. It should also serve as a testing ground or a lab for identifying challenges and where high-risk AI is meeting the populations who are fundamentally at risk, and who face so many social, cultural and economic challenges. AI technologies must be applied in ways that respect their dignity.

The second approach involves thinking about how to support these communities, not simply as users of AI or as recipients of policy makers views on how to protect them. It focuses on becoming direct architects of both products and policies. He asked that members of the panel and the audience to conceptualize how one builds policies that support not only the citizens of OECD countries but that actually seeks the input and participation of communities “at the front lines”.

The third approach that Mr Dhar discussed involved social institutions that are essential to making the promises of policy makers happen. This implies institutions that can coordinate and bring together various functions that cut across sectors. These applications, while they may not have direct market incentives, have direct responses and results in terms of impacts on populations across the world.

With each of these, at the Patrick J. McGovern Foundation and across civil society, everyone has the responsibility of becoming deeply digitally literate. This means understanding not only the challenges of AI that face the world today but to truly understand the opportunities that AI presents for humanity going forward. Finally, he expressed how enthusiastic the Foundation is about the OECD's approach to AI.

Ensuring that civil society is involved in solving AI's challenges

Mr Pilat thanked Mr Dhar and agreed with his perspective on how foundations and civil society can help to take things forward. He then asked Mr Dhar to tell attendees how to ensure that civil society has a seat at the table. He noted that they are typically present at the OECD, particularly in the work on digital economy policy.

Mr Dhar said that where he sees the most optimism is civil society's direct engagement in questions around privacy, dignity and making sure that commercial applications are built to support those who are most in need. But he continued by saying that there is a more fundamental issue. The vast majority of civil society is not technically oriented, meaning that they have yet to grapple with questions about AI. Their expertise lies in the social challenges that face people around the globe today. He expressed his belief that it is the responsibility of policy makers, the industry and corporate members that have spoken today and operate around the world, but also foundations like the Patrick J. McGovern Foundation, to bridge the gap. They should build the story of digital literacy that brings these organisations with their expertise and the institutes that they support to work in true participation and collaboration in fora just like the OECD.

AI GOVERNANCE AND SOFT LAW IN KOREA

Mr Pilat introduced Mr Haksoo Ko, Professor, Seoul National University Law School, Asia-Pacific Law Institute Director. He asked Mr Ko to discuss the state of governance and soft law surrounding AI in the Korean context and how it relates to other parts of the world from an academic and legislative perspective.

Mr Ko said that what has been happening in Korea has been very similar to what has been happening elsewhere in the sense that there have been so many discussions on many interrelated aspects of ethics. Over the past three years, government agencies have published a lot of ethics principles and guidelines. Here, the Ministry of ICT has played the leading role, along with a few other ministries and agencies. Some are general in nature, while others are sector-specific. For instance, there are guidelines for financial services, automated vehicles, personal data and others. Some private companies have published their own company guidelines.

Soft law or legislation?

In terms of hard law, several legislative proposals are pending but at the moment it is not entirely clear whether or not they have a chance of becoming law. It is also debatable whether or not it is beneficial to have hard law in place. The current debate around the EU's AI Act is getting a lot of attention in Korea. It is a multistakeholder issue that goes beyond government and industry. Civil society and the whole population need to be involved. There needs to be something between the part of society that wants technology to advance as fast as possible and parts that want time to discuss its pros and cons. A unique aspect of governance in Korea today is the Presidential Committee for the Fourth Industrial Revolution. It has two chairpersons. One represents the government and the other represents non-governmental stakeholders. The Presidential Committee has been serving as the main window of communication with all stakeholders.

Making Korea's AI framework interoperable

Mr Pilat then asked Mr Ko how to make sure that what is happening in Korea can be interoperable and comparable at the international level, with what is happening in other countries.

Mr Ko said that dialogue and communication are crucial and as the EU AI Act proposes, having a risk-based approach is a very good way of seeing AI technology, but context is very important. Countries have different attitudes towards risk and different calculations as to what is at stake and what risk means for their own society. Having some kind of referential between different stakeholders would be a good way to keep the dialogue open, but also among different countries and jurisdictions. The EU AI Act reflects its idea of risks, but Korea and other countries may have different assessment tools. He also pointed out that there are different definitions of fairness, depending on the country and the technology. Preparing referential and comparison tools could be a big step forward. He suggested that it could start at the OECD level and then expand for the whole world.

PRIORITIES FOR WORKERS, CITIZENS AND SMES IN THE AI TRANSITION

Mr Pilat welcomed Amanda Ballantyne, Director, Technology Institute of the American Federation of Labor and Congress of Industrial Organizations (AFL-CIO). He mentioned that AI is expected to change the world of work and asked her to talk about what she sees as priorities for workers, citizens and SMEs in the upcoming transition and the role that skills and upskilling need to play.

Ms Ballantyne began by saying that the impact of AI and digital technologies on the workforce is an open question. With proper regulation, AI could probably improve worker safety, support training and upscaling and increase productivity and allow workers to share in the benefits of productivity gains. But currently, with large technology firms based mostly in the United States and China, racing to develop AI systems without adequate regulation or the types of governance, democratic practices and transparency necessary for the creation of adequate regulation, we are at risk of exacerbating inequality polarization and discrimination as these technologies reshape our societies. The OECD AI Principles are an important milestone for steering these technologies in the right direction. Governments must now live up to their commitments to ensure that they are implemented.

She went on to say that including trade unions as a stakeholder in the labour market and policy implementation process is essential. They are not inherently opposed to the deployment of AI technologies. But they are concerned about how the technologies are used, who makes the decisions about their goals and who benefits from them. There are significant concerns about negative trends amongst member unions of the AFL-CIO, like bias and discrimination in AI human resource processes, increased worker surveillance and privacy issues, a lack of transparency in the design of workplace AI systems, unchecked automation and the dehumanisation of workers and degradation of work when monitoring systems are put in the workplace.

The dehumanizing effects of algorithmic management

She continued by saying that the dangers of dehumanizing algorithmic management are too clear in many of the stories that she has heard. For example, workers in Amazon warehouse facilities are directed by AI to achieve efficiency goals that injure their bodies. So much so that there are vending machines in the warehouses that only dispense painkillers. She then pointed out that mismanagement linked to algorithms are not just linked to workers in low wage sectors. Increasingly, unions that represent nurses and teachers are saying that their members must interact with patients and students in ways that are measurable by machines instead of using their own professional judgement, to deliver healthcare and education to vulnerable populations.

With the expansion of AI in management roles rather than assistance roles, we risk losing the invaluable role of human connection, compassion and judgment. Work can be hollowed out, middle-income professions are deskilled, and low wage workers have suffered increasing abuse and fragmentation. These and other examples show the need for government regulation in firm-level practices to curve the negative outcomes of AI implementation and this cannot be done without worker and trade union participation.

Everyone should have a say in how AI develops

She reminded attendees that the OECD AI Principles mention more than once and aim to ensure that the benefits of AI are broadly shared. From a trade union perspective, this effort has to be more than aspirational. She pointed out that for the last 70 years, enormous public investment has fostered the development of poor technologies required to develop and drive digitalization and AI technologies. But today we see deeply concerning market concentration among the leading AI firms and investors. According to a recent OECD report, the top ten companies filed for more than 20 per cent of AI patents between 2014 and 2018. The top fifteen companies own 14 per cent of all AI patents.

If AI companies are going to cause the transformations that proponents predict, the decision about what AI to develop and how to implement it cannot be left in the hands of an elite few. Workers have to be a part of the decision-making process at the firm and sector level through collective bargaining and also through their elected representatives in government.

She expressed the idea that everyone has a choice: AI can serve the agenda of a small group of elite players at the expense of workers, democratic values and the environment. Or, can we implement regulations that try to harness these technologies in the surface of promoting a more just, equitable and democratic world?

Inclusivity at the international level

Mr Pilat thanked her for her perspective on what needs to be done on the international level.

Ms Ballantyne stressed the importance of international cooperation when it comes to AI technologies it is very important for the OECD to foster cooperation amongst governments but also amongst civil society actors, with trade unions as a central actor in the talk about fair transition. She also suggested that the OECD and its partners identify three principles for the implementation of AI: choice, codesign and consultation. She went on to say that there is always a choice about whether or not AI should be implemented. She also emphasized the difference between AI as a facilitator for decision-making and AI as a decision-maker. She reminded attendees that the OECD AI Principles point this out as they discuss the connection between the development of human-centred AI. She then said that people need more collaboration around practices that ensure that AI is human-centric.

She continued by saying that the human-centred aspect of AI has to be rooted in codesign, development, public and private workers have to bargain over how and why AI systems are developed and deployed in sectors.

Finally, she evoked consultation, where the OECD has a major role to play fostering conversations among key actors on technological transformation in society including the deployment of AI. Since AI systems will impact employee performance, public investment decisions and the labour market, involving social partners including trade unions is essential.

QUESTIONS TO ALL PANELLISTS: WHAT SHOULD THE OECD'S MAIN PRIORITY BE?

Eric Horovitz

Eric Horovitz responded that there are twenty-some sets of principles issued by different organisations, including the OECD's. They are often great for high-level guidance but, he said, the devil is in the details. He recommended that each principle get down into the system and application levels with specific issues and concerns arriving in use cases and think deeply about the role of the human being in these systems and their role and influence in society and take it down to the next level on every principle, where "the rubber meets the road" technically and socio-technically. High-level principles are necessary and easy to formulate, and each suggests a pathway into deep discussion.

Andrea Renda

Andrea Renda echoed Mr Horovitz's advice and reiterated the need to consolidate, streamline differences and agree on a common path to better spread regulatory certainty to industry players, developers and deployers. He went on to say that the OECD has played an excellent role in sharing and exchanging best practices in participatory and experimental policy making. For this reason, he encouraged the OECD to create a platform for sharing sandbox results but also for discussing how to build the sandboxes in a way that makes sense, and where things are replicable, transferrable and scalable.

Vilas Dhar

Vilas Dhar also agreed with Mr Horovitz, adding that the principles are always the first step and that the challenge is figuring out what the next ten steps will be. He went on to encourage the OECD to conceptualize its role in building institutions that would hold the knowledge and the capacity to deliver on the promise of AI that lives outside of industry. This is about how to cite it geographically and socio-culturally in communities that need to be a part of this in the future.

Haksoo Ko

Haksoo Ko praised the OECD for providing a communications platform and accumulating cases and to continue to accumulate big cases related to AI ethics, as society demands for AI to be fairer and more transparent. Developers of AI systems should have some kind of quantifiable measure for transparency, fairness and other notions. Once a company has established its own matrix or experience, they can perhaps exchange those through the OECD from there everyone can start a conversation that could go into further details thanks to a database of use cases.

Andrea Renda

Amanda Ballantyne expressed agreement with what other panellists had said then stressed the importance of the OECD having broad and diverse stakeholders, and creating the capacity for that to happen to be able to conceptualize the totality of AI and not just its business implications. This means continuing the drive behind the conversation on fundamental rights, including key stakeholders who are not necessarily at the table in research and development AI and supporting the institutionalisation at the state level of the capacity and capability to convene stakeholders in developing the appropriate regulatory regimes for these technologies.

Vilas Dhar

Vilas Dhar also agreed with Mr Horovitz, adding that the principles are always the first step and that the challenge is figuring out what the next ten steps will be. He went on to encourage the OECD to conceptualize its role in building institutions that would hold the knowledge and the capacity to deliver on the promise of AI that lives outside of industry. This is about how to cite it geographically and socio-culturally in communities that need to be a part of this in the future.

OECD.AI

Policy Observatory