Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

10-Feb-2017

English text only

DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INNOVATION
COMMITTEE ON DIGITAL ECONOMY POLICY

DSTI/CDEP(2016)17
Unclassified

SUMMARY OF THE CDEP TECHNOLOGY FORESIGHT FORUM

ECONOMIC AND SOCIAL IMPLICATIONS OF ARTIFICIAL INTELLIGENCE

OECD Conference Centre, Paris, 17 November 2016

A half-day OECD Technology Foresight Forum was held on Thursday 17 November 2016, with speakers invited to present developments around Artificial Intelligence to the CDEP. The present document contains a summary of the presentations and has been reviewed by the invited speakers.

The agenda, presentations, bio book and final summary are available at: http://oe.cd/ai2016

Contact: Karine Perset: Tel: +33-1 45 24 93 54: e-mail: karine.perset@oecd.org
Anne Carblanc: Tel: +33-1 45 24 93 34; e-mail: anne.carblanc@oecd.org

JT03408833

English text only

TECHNOLOGY FORESIGHT FORUM 2016

THE ECONOMIC AND SOCIAL IMPLICATIONS OF ARTIFICIAL INTELLIGENCE SUMMARY

PRESENTATIONS AND DISCUSSIONS – MAIN POINTS

AI is intelligence exhibited by machines

Participants at the Forum defined Artificial Intelligence (AI) as the capability of a computer programme to perform functions usually associated with intelligence in human beings, such as learning, understanding, reasoning and interacting, in other words to "do the right thing at the right time". For example, machines understanding human speech, competing in strategic game systems, driving cars autonomously or interpreting complex data are currently considered to be AI applications.

Machine learning, big data and cloud computing have enabled AI's recently accelerated progress

Participants underlined that despite fluctuations in public awareness, AI has been making steady progress since its inception in the 1950s. They said that more recently, the availability of big data, cloud computing and recent breakthroughs in an AI technology called machine learning, have dramatically increased the power, availability and impact of AI. For example, Google's AI learns how to translate content into different languages based on translated documents that are online and Facebook learns how to identify people on images based on its existing large database of known users. In a short time, AI has become main stream and is rapidly permeating and transforming entire sectors of the economy and as a result, raises policy issues across numerous economic and social dimensions.

There are few limits to what AI will be able to do

Participants agreed that advanced Artificial Intelligence is already here and that there are few limits to what it will be able to do. There was a call to focus on 'Applied AI' that is designed to accomplish a specific problem-solving or reasoning task. However, several participants felt that policy-makers could not ignore the possibility of a (hypothetical) "artificial general intelligence" (AGI) whereby machines would become capable of general intelligent action, like a human being. Others cautioned that discussions should be realistic in terms of time scales and highlighted that AI, like natural intelligence, is necessarily constrained by what computer scientists term *combinatorics* – the inconceivably vast number of things an intelligent system might think or do. In addition, because AI is an artefact, AI systems are constructed using architectures that limit AI to the knowledge and potential actions that make sense for a given application.

AI promises to improve efficiency and to help people solve complex problems

Participants agreed that AI can improve efficiency, save costs and enable better/more accurate decisions, since it allows data processing at enormous scales and accelerates the discovery of patterns. They also provided examples of areas where AI enables whole new functionalities or new solutions. Some participants characterized most AI applications today as automating the detection of patterns in speech, images, videos or user behaviour. They gave examples of AI applications in basically every sector of the economy, including social networking, health, the environment, or transportation. For example, with IBM's Watson, AI systems that are designed to "augment human intelligence" help physicians to recommend courses of treatment. AI applications that are already in widespread use today but are not obvious include credit card payment checking, GPS navigation systems, search engines, translators, spell and grammar checkers, spam filters, social networks, personal assistants like Apple's SIRI, or robotic applications like vacuum cleaner robots. Discussants at the Forum were confident that AI would help people to solve complex global challenges such as those related to the environment, transportation or health by helping scientists to spot complex cause and effect relationships.

While some initiatives are underway, more AI awareness building is needed

AI is increasingly a priority for OECD Member and Partner countries, including the Japanese Ministry of Information and Communications (MIC) and Cabinet Office, the G7 ICT ministers meeting in April 2016, the U.S. White House or the U.K. Parliament, as well as for the private sector with for example a recently launched "Partnership on artificial intelligence to benefit people and society".

Participants at the Forum warned that overall however, the likely impact of AI in the years ahead was still underappreciated by policymakers and by the public at large. They foresaw that AI would transform entire sectors of the economy and lead to in-depth societal changes. A recurrent issue was the need for public policy makers from governments and intergovernmental organizations to understand and begin to prepare societies for the changes afoot. Policy makers were urged to consider how societies could best benefit from AI while minimizing risks of discrimination, privacy erosion, loss of public anonymity, as well as possibly setting limits for automated decision-making. There was also a call to replace AI sensationalism with evidence-based and informed communications on AI.

AI is poised to accentuate income inequality and impact human labour

One of the recurrent themes that emerged from the Forum was that of the potential unequal impacts of AI on income distribution and of who will control AI technology. On the economic side, the notion of "winner-takes-all" technology was highlighted as important. A few technology companies and governments with access to large amounts of data and funding could control AI technology, have access to its super-human intelligence and gather most of the AI benefits. AI means that companies will rely less on human workforce, meaning that revenues will go to fewer people. One speaker brought up the possibility of a universal basic income. There were also concerns about the geopolitical impacts of this winner-takes-all paradigm, with AI likely to affect labour from developing countries disproportionately, while developed countries largely control the technology, patents and data. Some companies seem to be taking action. A participant from Google explained that Google distributes its machine learning technology freely via open source, enabling others to benefit from it.

Projections of the impact of AI on jobs were extremely variable. AI was expected to augment or replace human labour not just in manual work but also in highly-educated professions such as radiology. Participants said that jobs at risk of being replaced by AI tended to be predictable or repetitive jobs such as drivers, cashiers, CPAs, factory workers, or doctors. Some anticipated creative destruction and the creation

of entirely new types of job. One participant emphasized the need for people to find meaning in other, non-labour, activities and to learn new ways to contribute to human society.

Reasonable transparency, or explainability, of AI

One of the strongest themes that emerged from the Forum was the need for transparency of AI algorithms. AI algorithms that make decisions impacting people and/or that handle critical infrastructure were identified as requiring particular transparency. Participants stressed the need to be able to find out which inferences AI makes on people and to question the validity of such inferences and of the resulting decisions. The discussion also explored the limits of transparency when algorithms cannot necessarily be understood by humans. The concept of "reasonable transparency" was put forward. Others noted that "explainability" or "understandability" may become more useful concepts than transparency, as machine code generated by AI rapidly becomes too complex for humans to decode. Governments were viewed as having a particular role to play in ensuring transparency or explainability of AI. They widened the discussion to trust in AI, which would also require transferability and traceability of AI.

Designing human-centric AI, that respects the rules of societies

Another theme that emerged from the Forum was that of "ethics by design", as discussants pointed out risks such as the ranking of people or their manipulation by AI based on knowledge of their preferences. Some participants also pointed out that AI reflects both the knowledge and the prejudices embodied in data. AI researchers put forth that AI systems could be designed to integrate "ethics by design" respecting people from the start of the systems design: researchers from commercial entities that own or create AI would need to design agents' specifications that respect the rules that societies have created for themselves, respect cultural and religious differences and are appropriate for the use contexts (*e.g.* professional, home, etc.). There was, however, a call for more dialogue on how a concept like "AI Ethics" or "Ethics by design" might be applied and moreover, applied at a transnational level, since both ethics and the laws codifying ethics are often national.

One participant noted that while Europeans tend to focus AI ethics discussions on risks to privacy and fundamental human rights, other parts of the world such as North America focus, for example, on discrimination. A speaker cautioned that AI in the wrong hands could increase the risks to human rights and civil liberties with applications such as predictive policing and facial recognition. Several participants also contrasted the different research and implementation approaches taken to AI in different regions of the world, notably contrasting the individual-centred cognitive approach being taken towards AI in the West to the society-centred connectionism adopted in the East. Several speakers emphasized the need, everywhere, to pay more attention to AI's predictive power.

Responsibility and liability, security and safety

Participants highlighted how AI-driven automated decision-making raise questions of responsibility and liability, for example when life-threatening accidents involve driverless cars. They emphasized the "authored"/machine nature of AI and the fact that AI cannot usefully be made a legal person, or hold responsibility for its decisions. Once again, the example was given of Google which had said it would be liable for its autonomous vehicles. Participants also highlighted emerging AI safety and security risks, for example the risk that malware could abuse an AI network system or that autonomous weapons could be used maliciously.

The need to invest in human skills to leverage AI

Speakers called for policy adjustments to facilitate continuous education, training and skills development, with more AI theory and practice at university. They expected that significant adaptation and

adjustments would be needed as people found how best to participate in society and leverage their humanity, individuality and creativity in a world where AI may obviate the need for many existing human jobs. Initiatives such as IBM's Dr Watson position AI as supporting human decisions rather than replacing humans, by enhancing and scaling human expertise.

Articulating common principles for AI would be timely

Finally, the roles of different stakeholders were emphasized. The key role of governments and intergovernmental organisations was stressed in areas such as setting or codifying rules of the game, in respect to societal and economic impacts, planning for the long-term and international co-operation. Several participants said that policy makers should convene stakeholders to consider social, economic, ethical, and legal issues of AI networking. They encouraged the OECD to develop general AI guidelines. Some speakers felt the need for international guidelines, along with culture-appropriate national laws, even if this increased time-to-market. Others added that critical or dangerous areas could be differentiated and could for example trigger human intervention. Some participants suggested that it may even be necessary to prohibit some applications of AI research or application, for example autonomous AI killer weapons. AI that simulates emotions, or "synthetic AI", was also viewed as another area that may require safeguards and particular transparency.

PRESENTATIONS – SUMMARY

Mr. Jörgen Abild Andersen, CDEP Chair, introduced the issues and speakers and moderated the discussions.

1. Welcome Remarks

Mr. Andrew Wyckoff, Director for Science, Technology and Innovation, OECD, explained that every 2 years delegates in the OECD Committee on Digital Economy Policy (CDEP) choose an issue that they feel raises particular opportunities and challenges for policymakers: this year CDEP chose Artificial Intelligence (AI). He highlighted that AI is increasingly a priority for OECD Member and Partner countries, including the Japanese Ministry of Information and Communications (MIC) and Cabinet Office, the G7 ICT ministers meeting in April 2016 or the U.S. White House, as well as for the private sector with for example the recently launched "partnership on artificial intelligence to benefit people and society". He welcomed speakers from research and academia, business and the public sector and thanked the Japanese Ministry of Information and Communication for its support and for encouraging the OECD to look at public policy considerations associated with AI.

2. Key Artificial Intelligence (AI) Developments and applications: Opportunities and challenges today and tomorrow

Mr. Yves Demazeau, Research Director, Centre National de la Recherche Scientifique (CNRS), France defined Artificial Intelligence (AI) as the imitation of human behaviour using informatics and electronics. He said that AI is concerned with knowledge rather than just data and foresaw that AI and data processing (aka «Data Science») would merge. He highlighted that while AI research has steadily progressed over the past 60 years, the media had recently focused on Machine Learning/Deep Learning as a very visible part of automatic learning. He presented key AI application areas for *i)* Everyday life (*e.g.* intelligent houses, energy management, language translation, personal assistants, support to people), that raise privacy and data protection issues; *ii)* Transportation (*e.g.* transport networks, smart cities, autonomous vehicles like subways, cars or planes), that raise legal/liability issues; and, *iii)* Society (*e.g.* interactive games, social networks or matchmaking, news personalisation) that could raise moral issues. He stressed the likely implications for data protection and privacy of AI penetrating all sectors, from agronomy or virtual reality to medicine or digital humanities.

He provided an overview of regional AI developments, contrasting the individual-centred cognitive approach taken towards AI in the West to the society-centred connectionism adopted in the East. He mentioned the Chinese government's recently launched AI initiative and said that large Chinese technology companies (*e.g.* Baidu or Alibaba) are investing significantly in AI, as are large American companies (*e.g.* Google or Amazon). He recalled the leadership of Japan in AI, with the MIC's early involvement, the Japanese government's national plan that includes AI, large investments by Japanese technology and robotics companies such as Fujitsu, Hitachi, NEC, as well as joint labs between universities and companies. In his view, the EU and France are lagging.

Turning to the future of AI, he emphasized the rapid growth of artificial agents and their utility to solve complex problems. He stressed the need to build awareness of policy makers distinguishing facts from opinions, to invest in all AI disciplines rather than just deep learning, to improve education on AI at

universities, and to plan for ethics by design i.e. adapting agents' specifications to the use context and respecting cultural and religious differences.

Mr. Olivier Ezratty, Innovation Strategies Consulting, said that most existing AI applications automate pattern detection in speech, images, videos or user behaviour. He attributed the progress in pattern detection to significant progress in IoT networks, lots of data, and to the availability of computing power through the cloud. He provided an overview of key AI startups in the areas of deep learning (*e.g.* Numenta, Sentient); predictive analysis (*e.g.* Context Relevant, Work Fusion); visual search (*e.g.* Vicarious, Clarifai), chatbots (*e.g.* Existor, Msg.ai), and vertical apps and emphasized the potential of AI for healthcare, transportation or the environment.

M. Ezratty reported extremely variable predictions of the impact of AI on jobs and pointed out that jobs at risk of being replaced by AI would likely be repetitive (*e.g.* cashiers, CPAs, factory workers, drivers, doctors). He also stressed the importance of raising awareness of policy makers; of public AI investments in key areas like healthcare, environment /energy, transportation; of adapting university education and of fostering innovation.

Mr. Jonathan Sage, Government and Regulatory Affairs Executive, EMEA lead on cyber security and cloud computing policy, IBM, presented AI as the capability of a computer program to perform tasks usually associated with intelligence in human beings. He said that IBM refers to "Augmented Intelligence", since AI enhances and scales human expertise and decision-making, rather than replacing humans and their judgment, emotions, instinct, etc. Cognitive Computing at IBM includes machine learning, reasoning and decision technologies, language, speech and vision recognition and processing technologies as well as human interface technologies.

He gave examples of existing AI applications, including: credit card payment processing, GPS navigation systems, spam filters, social networks, personal assistants like Apple's SIRI, and robotic applications like vacuum cleaner robots. He provided more information on IBM's Watson, an AI platform with applications for example in healthcare, whereby systems help physicians to recommend courses of treatment. He noted other promising applications of Watson such as for cybersecurity that requires the processing of huge amounts of unstructured data into knowledge. He highlight IBM's AI policy priorities: 1) privacy; 2) safety and security; 3) algorithmic transparency and 4) assessing impact on jobs.

Ms. Ophélie Gerullis, Public Policy Manager, Facebook presented the major AI applications at Facebook and work being conducted by the company's researchers. She said that AI is a top priority for Facebook as it designs systems that can perceive and reason about their environment through: i) Language: speech recognition, language translation, question answering, dialog system (*e.g.* using chatbot messenging to book a taxi or schedule a haircut); ii) Vision: face detection/recognition, object/text recognition, action classification, natural language description, etc.; and, iii) Planning: given a starting point and an end goal, planning strategy. She noted that challenges remain in applications like image recognition that are highly contextual and with strategy planning.

Ms. Lynette Webb, Senior Manager, European Policy Strategy, Google, said that machine learning is already a key part of most of Google's products and services: it hugely boosts quality and accuracy and enables new functionalities. For example, Google Translate uses machine learning for 8 language pairs: rather than words, full sentences are translated with much better results. She showed an example of a mobile translate app that automatically translates and displays text from a picture taken with a camera phone. She added that Google also leverages AI to detect spam and phishing, to filter out pornography/violent imagery with SafeSearch and to allow YouTube channel owners to review potentially offensive comments.

In terms of policy, Ms. Webb said that Google aims to open its AI / machine learning and enable others to benefit from it. Google TensorFlow is available freely on an open source basis and Google also commercializes Google machine learning on the Google Cloud Platform via both pre-trained models (a Vision API to understand images, a Natural Language API to analyse text and a Translate API to convert text into another language) and customizable models. For Ms. Webb, we are just starting to apply AI to real world problems; spotting patterns that humans cannot and helping scientists understand hugely complex cause and effect relationships. Research challenges remain notably to improve the understandability of models as there is currently a trade-off between accuracy and understandability. She pointed to open implementation issues notably ensuring that systems are fair/functioning, determining whether there are areas in which AI should not be used, and supporting the transition in jobs/employment.

3. AI and society 5.0

Ms. Yuko Harayama, Executive Member, Council for Science and Technology Policy, Cabinet Office of Japan (via video link from Tokyo) provided an overview of AI policy developments in Japan. She explained that within the Cabinet Office of Japan, the Council for Science, Technology and Innovation (CSTI) had created an Advisory Board on AI and Human Society in May 2016. The Advisory Board is working to prepare a human-centred "Society 5.0" that combines sustainability, inclusiveness, efficiency and intellect. She emphasized the social dialogue needed to address AI issues of: Ethics (*e.g.* manipulation by AI, preferences, ranking of humans), Law (*e.g.* privacy, responsibility for accidents with self-driving cars), Economy (*e.g.* benefiting from AI while addressing AI income inequality), Society (*e.g.* dependence on AI, "off switches" to allow people to not to use AI), Education (*e.g.* exploiting human strengths, ability to exploit AI), and R&D (*e.g.* to ensure AI is secure, transparent, controllable, and ethical).

She emphasized the social responsibility and moral imperative we have to address policy challenges in order to benefit from AI (*e.g.* with low cost personalized services) while minimizing risks (*e.g.* privacy, discrimination, loss of public anonymity etc.) and setting the limits of automated decision making.

4. Public policy considerations raised by AI

Mr. Susumu Hirano, Professor, Faculty of Policy Studies, Dean, Graduate School of Policy Studies, Chuo University focused on possible safeguards to mitigate the risks of AI at the R&D stage and on transparency, safety, ethics, accountability and social acceptability. Recalling the studies conducted by the Japanese MIC, he foresaw progress of AI networking: 1) stand-alone AI that functions independently from the Internet; 2) formation of networks of AI that coordinate autonomously and harmonize their progress in various sectors of society; 3) AI extends humans' physical and intellectual capabilities, and; 4) finally, co-existence of humans and the AI network system, with information in human brains being output externally and, for example, robots being operated by human thought.

He recalled that at the G7 ICT Ministers' Meeting in Takamatsu, Japan (April 2016), participating countries agreed with the proposal by Minister Sanae TAKAICHI, Japanese Ministry for Internal Affairs and Communications (MIC), to i) convene stakeholders to consider social, economic, ethical, and legal issues of AI networking, and ii) formulate principles for AI development. 8 draft principles modelled on the OECD privacy guidelines were proposed for AI network systems: 1) transparency, 2) user assistance, 3) controllability, 4) security, 5) safety, 6) privacy, 7) ethics, and 8) accountability. He applied the draft Principles to the (hypothetical) bridge case, whereby an AI driverless car must make an ethical choice between minimizing casualties or protecting its occupants. As the next step, MIC will host an International Forum toward AI Network Society in Tokyo in March 2017 to discuss, inter alia, progress of the Subcommittee on AI R&D Principles of the 'Conference toward AI Network Society' and to advance international discussions of AI and AI R&D Guidelines.

Dr. Joanna Bryson, Reader at University of Bath, and Affiliate, Center for Information Technology Policy at Princeton University cautioned that advanced artificial intelligence is already here and that there are few limits to what it will be able to do. She defined "intelligence as doing the right thing at the right time" (in a dynamic environment), and said that it requires: 1) Detecting the context, 2) Learning the actions that can be performed and 3) Associations between these. She characterised Artificial General Intelligence as a myth in that, like natural intelligence, AI is constrained by both the mathematics of combinatorics, and by the people who construct architectures limiting AI to actions that make sense. Stressing the difficulty of search if all possible options must be investigated (combinatorics), she attributed human intelligence to humans' skill at communicating solutions once found, and – more recently – at mining these solutions from our culture as input to machines. Machine learning therefore exploits existing search to find the right thing to do at the right time.

She pointed out that being neither human nor moral; AI reflects both the knowledge and prejudices embodied in data. Defining ethics as the way a society defines and governs itself, she noted that determining AI's place in society is a normative decision and not a fact for science to discover. She recommended the UK's EPSRC Principles of Robotics that assert that robots are ordinary products that should: *i)* not be designed as weapons; *ii)* be designed and operated to comply with existing law including on privacy; *iii)* be designed to be safe and secure; *iv)* as manufactured artefacts, should not use the illusion of human-likeness to exploit vulnerable users; and *v)* should be licensed with clear ownership by a legal person. She added that efforts to create legal personhood for robots might allow companies to choose to utilise robots to escape legal and financial liability and encourage tax avoidance. She further noted that AI and ICTs more generally make the entire world essentially employees of companies that are taxed in at most one country – often not even well there because the exchanges (both labour and payment) are of information which is not well denominated.

She questioned that AI would create unemployment. In her view, AI reduces human differentiation and makes it possible to hire less skilled employees. In addition, those who earn money tend to spend it, creating new forms of employment. She noted that people already have robots in the form of smart phones and are connecting their governments, companies and families into the same space. In her view, the largest AI issue will be that of redistribution as inequality worsens, and of supporting human differentiation through privacy.

Ms. Cornelia Kutterer, Director of Digital Policy, Microsoft EMEA noted the key role of the cloud in making computing resources widely available and said AI is one component of the digital transformation of economies and societies. She said that AI provides opportunities to augment human capabilities, improve society, enable vast amounts of data, AI affordable to all and capture global economic benefits. The challenges associated with AI include the societal change/employment, access to data / inequality, ethical concerns.

A policy priority in her view is to modernize traditional laws holistically in the areas of data access, copyright, trade secrets, safety and liability, transparency, and publicly available data. She also thought privacy practices should evolve to facilitate innovation and data repurposing. In her view, Europeans tend to focus AI ethics discussions on privacy and fundamental human rights, while other parts of the world focus, for example, on discrimination. In her view, more attention to AI's predictive power is needed. Ethical practices require transparency, non-discrimination, multi-stakeholder collaboration, industry standards, and cloud-powered AI. Citing Microsoft's CEO Satya Nadella, she said that AI should 1) be designed to assist humanity, 2) be transparent, 3) maximize efficiencies without destroying the dignity of people, 4) be designed for privacy; 5) have algorithmic accountability so that humans can undo unintended harm, 6) guard against bias. She highlighted initiatives that Microsoft is involved in; notably the AI Partnership, the Ethics of Medical data analytics, and the IEEE work on Advancing technology for the benefit of humanity

Mr. Cyrus Hodes, Director for Artificial Intelligence and Representative for the MENA Region, The Future Society @ Harvard Kennedy School of Government focused on the gap between practitioners of AI and Policy makers, on the types of legal and technological safeguards needed for AI and on ideas on how to make this happen. He highlighted risks AI poses to: employment, social welfare and equality; privacy, agency and dignity; security and even humanity. Regarding employment, he questioned whether AI would lead to a Schumpeterian wave of creative destruction or to great unemployment, with AI causing both medium and highly-skilled jobs to be replaced with low-skill jobs. He contrasted the approach of designing AI that augments or enhances humans to the approach of designing intelligent 'agents' that would be able to learn, adapt and deploy in dynamic and uncertain environments.

He stressed the importance of public policy makers beginning a dialogue on the implications of AI in fora such as the OECD, the EU or the UN. He said that for both the private sector and governments a global race is already underway to acquire AI in winner-takes-all and first-mover-advantage paradigms. Policy-makers need to engage in: 1) the revolution in education and training (*e.g.* how humans labour can complement AI, deploying accelerated professional transitions) and 2) redistribution mechanisms from AI winners, concentrated in a few private sector firms, to AI losers, stressing the need for a global dialogue and possible need to consider universal income. He emphasised the importance of global coordination and of governments convening defence industries together to discuss ethical considerations of killer robots and drones.

Mr. Tatsuya Kurosaka, Project Assistant Professor at Keio University Graduate School of Media and Governance spoke of the effort in Japan to anticipate the social and economic impacts of AI networking from 2040 to 2040. He said that assessing future impacts on different industries is difficult due to rapid evolution and innovations, a Study group had attempted to evaluate the impact of AI on different industries, such as public infrastructure, public administration, personal assistance or industries such as transportation/logistics, finance /insurance, or medical care/nursing. One of the areas is that of AI disaster prevention applications, important for Japan, that help manage natural disasters and lessen their damage. In the area of finance and insurance for example, the automation of trading, loan screening, and credit management is expected to become widespread around 2030.

The conference conducted a risk scenario analysis after classifying social, economic, ethical, and legal risks caused by AI networking, notably risks of: 1) functions in the AI network system not working appropriately; 2) an AI network system infringing rights or interests and causing an accident, a criminal attack or an attack by an autonomous weapon system; 3) infringement of privacy and personal information, *e.g.* with AI surmising people's intentions, health, or future actions; 4) human dignity and the autonomy of individuals with *e.g.* invisible manipulation of human decision-making processes and; 5) democracy and governance mechanisms, whereby decision-making processes become opaque and responsibilities ambiguous. Taking into account the unpredictability of AI developments, the work focused on reasonable transparency that users and experts could work together on.

5. Wrap-up and next steps

Ms. Anne Carblanc, Head of OECD Division on Digital Economy Policy thanked the speakers and summarised the Forum's discussions. She highlighted that we are just beginning to understand the impact of AI and the challenges that it raises in areas such as employment and redistribution; security, privacy, responsibility, transparency; as well as training and continuing education. She noted that there was a consensus on the need for a human-centric AI. She indicated that the next steps would be for the OECD to conduct analytical work on AI in 2017 and convene a multi-stakeholder discussion in September 2017. Then, if CDEP Members would agree, the OECD could consider developing guidance or a draft Council Recommendation on AI during the course of 2018, building on work underway in some countries like Japan, the United States or the United Kingdom, as well as private sector or research initiatives such as the AI Partnership.[1]

[1] http://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/news-parliament-2015/robotics-ai-report-published-16-17/

APPENDIX A: AGENDA AND SPEAKERS

Chair of the Technology Foresight Forum: Mr. Jörgen Abild Andersen, CDEP Chair

1. WELCOME REMARKS

Mr. Andrew Wyckoff, STI Director, OECD

2. AI DEVELOPMENTS KEY APPLICATIONS AND ECONOMIC IMPACT OF ARTIFICIAL INTELLIGENCE (AI) : TODAY AND TOMORROW

Mr. Yves Demazeau, Research Director, Centre national de la recherche scientifique (CNRS), France

Mr. Olivier Ezratty, Innovation Strategies Consulting

Ms. Cornelia Kutterer, Director for Digital Single Market Strategy, Microsoft Corporation

Mr. Jonathan Sage, AI, Cloud Computing and Cyber Security policy lead, IBM Europe

Ms. Ophélie Gerullis, Public Policy Manager, Facebook

COFFEE BREAK

3. AI AND SOCIETY 5.0

Ms. Yuko Harayama, Executive Member, Council for Science and Technology Policy, Cabinet Office of Japan *(via video link)*

4. PUBLIC POLICY CONSIDERATIONS RAISED BY AI

Mr. Susumu HIRANO, Professor, Faculty of Policy Studies, Dean, Graduate School of Policy Studies, Chuo University

Ms. Joanna Bryson, Reader at University of Bath, and Affiliate, Center for Information Technology Policy at Princeton University

Mr. Cyrus Hodes, Director for Artificial Intelligence and Representative for the MENA Region, The Future Society @ Harvard Kennedy School of Government

Mr. Tatsuya Kurosaka, Project Assistant Professor at Keio University Graduate School of Media and Governance

5. WRAP-UP AND NEXT STEPS

Ms. Anne Carblanc, Head of OECD Division on Digital Economy Policy

Up-to-date information at the following address: http://oe.cd/ai2016