SNU AI Policy Initiative
http://ai.re.kr/

서울대학교 법과경제연구센터
SNU Center for Law & Economics

2019 Conference Report

AI Policy for the Future: Can We Trust AI?

Co-directors **Haksoo Ko and Yong Lim**
Report prepared by **Do Hyun Park**

Date/Time : Friday, **August 23,** 2019 09:00am-17:00pm

Venue : Korea Press Center International Conference Hall (20F)

Host : 서울대학교 법과경제연구센터 SNU Center for Law & Economics

Sponsors : NAVER Google aws

Conference videos available on:
Naver Channel https://tv.naver.com/v/10550957/list/524105
YouTube https://www.youtube.com/playlist?list=PLGEVmkdHcJTKcgVvUhEcCDDiHKDUY6MSn

SNU AI Policy Initiative
http://ai.re.kr/

2019 Conference Report

AI Policy for the Future: Can We Trust AI?

Co-directors **Haksoo Ko and Yong Lim**
Report prepared by **Do Hyun Park**

Date/Time : Friday, **August 23,** 2019 09:00am-17:00pm

Venue : Korea Press Center International Conference Hall (20F)

Host : 서울대학교 법과경제연구센터 SNU Center for Law & Economics

Sponsors : NAVER Google aws

Conference videos available on:

Naver Channel https://tv.naver.com/v/10550957/list/524105

YouTube https://www.youtube.com/playlist?list=PLGEVmkdHcJTKcgVvUhEcCDDiHKDUY6MSn

속표지

Seoul National University

SNU AI Policy Initiative

(http://ai.re.kr/)

Co-directors : Haksoo Ko and Yong Lim

Report prepared by: Do Hyun Park

(This report is a translation from the original conference report which was prepared in the Korean language.)

Conference Report

AI Policy for the Future: Can We Trust AI?

Date: August 23, 2019 (Friday), 9:00–17:00

Venue: Korea Press Center International Conference Hall (20F)

Host: Seoul National University Center for Law and Economics

Sponsors: Naver, Google, Amazon Web Services

Conference videos available on:

Naver Channel : https://tv.naver.com/v/10550957/list/524105

YouTube : https://www.youtube.com/playlist?list=PLGEVmkdHcJTKcgVvUhEcCDDiHKDUY6MSn

Conference Program

Master of Ceremonies : Kyoungjin Choi (Gachon University)

Registration 9:00–9:30

Welcoming Remarks 9:30–9:35

- Haksoo Ko (SNU AI Policy Initiative Co-director)

Congratulatory Message 9:35–9:45

- Wonki Min (2nd Vice-Minister of Science and ICT)

Session I 9:50–12:40

Governance of AI: What Needs to be Done to Achieve Public Trust in AI-impacted Outcomes?
Moderator: Yong Lim (SNU AI Policy Initiative Co-director)

Keynote Speech I 9:50–10:15

- R. David Edelman (MIT)

Keynote Speech II 10:15–10:40

- Deirdre Mulligan (UC Berkeley School of Information)

Panel Discussion 11:00–12:40

Panel
- R. David Edelman (MIT)
- Hyunseop Kim (Seoul National University)
- Jake Lucchi (Google)
- Deirdre Mulligan (UC Berkeley School of Information)
- Kyung Hee Song (Director General, Ministry of Science and ICT)
- Jeongwon Yoon (Amazon Web Services)

Session Ⅱ 14:00–16:50

Fairness in AI: What Does It Mean, and How Can It be Implemented?
Moderator: Yong Lim (SNU AI Policy Initiative Co-director)

Keynote Speech Ⅲ 14:00–14:25

- Fredrik Heintz (Link ping University, Sweden and EU High-Level Expert Group on Artificial Intelligence)

Keynote Speech Ⅳ 14:25–14:50

- Blaise Agüera y Arcas (Google AI)

Panel Discussion 15:10–16:50

Panel
- Norberto Andrade (Facebook)
- Blaise Agüera y Arcas (Google AI)
- Gary Chan (Singapore Management University)
- Fredrik Heintz (Linköping University and EU High-Level Expert Group on Artificial Intelligence)
- Malavika Jayaram (Digital Asia Hub, Hong Kong)
- Indrė Žliobaitė (University of Helsinki, Finland)

Wrap up and Closing 16:50–17:00

International Conference
AI Policy for the Future: Can We Trust AI?

Introduction

Can we trust artificial intelligence (AI)? This is a timely question in these days when we witness people dying of self-driving car test runs and being attacked by drones.[1] Given that the society's trust is closely associated with acceptance of AI technologies, some see that appropriate responses to such a question will be a decisive factor for the success or failure of the AI industry.[2] Against this backdrop, the Seoul National University (SNU) Center for Law and Economics initiated the annual conference in 2017, where experts are invited to discuss legal and policy issues surrounding big data and AI technologies.[3] From the second conference in 2018, a conference report like this was produced as part of its multi-faceted efforts to maximize the effects of the conference.[4]

---

1) Recently, the U.S. National Transportation Safety Board reached a conclusion that in the Tesla autopilot vehicle-truck crash last year, the autopilot system was at least partially responsible for the accident. Meanwhile, the recent drone attack to the Saudi Arabian refinery resulted in a surge in international oil prices and spilled over to social problems.

2) High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", European Commission (2019. 4. 8), pp. 4-5.

3) Entitled "Policy Issues surrounding AI, Algorithms and Privacy", the first conference was centered around discussions over policy tasks surrounding AI, algorithms, and privacy in the opening session, followed by individual sessions that focused on "AI big data and market competition", "decision making by AI and legal and social accountability", and "data de-identification". The KDI filmed the first conference and uploaded the video clips, which are available at the following URL.
http://www.youtube.com/playlist?list=PLOP6ilKzhDLQ_a2hMmD0vxsJn0d-aQco8.
The second conference was entitled "Artificial Intelligence Today: Governance and Accountability". The focus of the first session was on how to build a right data governance, and the second and third sessions were centered around discussions over accountability and ethics related to automated decision-making, and how to regulate blockchain and other new technologies, respectively. SNU undeclared majors students Ji Hyun Lee, Jae Seung Park, Keon Hee Yoon, and Young Chae Cho filmed the second conference and uploaded the videos, which are available at the following URL.
http://tv.naver.com/aipolicyintiative; http://www.youtube.com/channel/UCKyxSZOtLB1YvkKM2_Mq8gQ.

4) 2018 Conference Report is available at the following URL.
http://ai.re.kr/%ea%b3%b5%ec%a7%80%ec%82%ac%ed%95%ad/?lang=en

The SNU Center for Law and Economics established the SNU AI Policy Initiative in 2017 as a research program to look into and discuss social, economic, legal, and ethical issues following the advancement and increasing use of the AI technology and their policy implications.

The overarching topic of the third AI conference in 2019 was how to achieve trust and fairness in the era of AI. Trust and fairness are two closely-associated concepts as earning people's trust in AI requires aligning the ripple effect brought by AI with the humanity's value system of fairness. The first session was centered around the former, discussing governance structure for AI to earn public "trust." The second session focused on the latter, i.e., discussions over what "fairness" in AI means and how to achieve it. A notable point of this year's conference was two keynote speeches preceded discussions in each session with the aim to further vitalize discussions with insightful views presented by authorities in relevant fields.

The involvement of renowned figures from home and abroad furthered the depth of discussions at this conference. In particular, Second Vice-Minister of Science and ICT Wonki Min who delivered the congratulatory remarks is an AI expert and served as the chairman of the AI Expert Group at the OECD. He is recognized as one of the key contributors to the development of the OECD AI principles. All other keynote speakers are some of the most authoritative figures in their respective fields, whether it be academia or industry. Academic leaders in AI delivered keynote speeches for the first session. Professor R. David Edelman from the Massachusetts Institute of Technology is acknowledged as an authority with experience in both industry and academia and drew up IT policies for the White House for years. Professor Deirdre Mulligan from the University of California at Berkeley is a renowned scholar with remarkable achievements in studies of privacy and fairness. The keynote speakers for the second session were the current leaders in AI practice. Professor Fredrik Heintz from Linköping University is a member of the EU High-Level Expert Group on Artificial Intelligence, and Blaise Agüera y Arcas is a chief scientist and engineer that leads Google's AI group.

In addition, efforts were made to ensure inter-disciplinary convergent thinking beyond the boundaries of technology and law by inviting prominent and young scholars from Americas, Europe, and Asia and involving Professor Hyunseop Kim who teaches philosophy at SNU. Furthermore, careful considerations were given to diversify the backgrounds of the invited speakers so as to give attention to the voice of governments and businesses rather than merely focusing on academic views. For example, in addition to the Vice-Minister Wonki Min mentioned earlier, public officials such as Director General Kyung Hee Song from the Ministry of Science and ICT and incumbents working for global enterprises such as Google, Amazon, and Facebook were invited with the aim to ensure this conference would reflect viewpoints from many different fields and areas. Such an aim was fulfilled by the presence of more than 400 people attending the conference, who came from many different backgrounds ranging from law to engineering, economics, philosophy, etc.

Ⅰ. Session 1 - Governance of AI: What Needs to be Done to Achieve Public Trust in AI-Impacted Outcomes?

1. Keynote Speech Ⅰ
– Governing Artificial Intelligence: How Machines Can Earn Our Trust?

Professor Edelman started his keynote speech by touching on the topic of "disruptive innovation."[5] Indeed, media have relied on this term to present rosy outlooks for the outcomes of the so called Fourth Industrial Revolution represented by AI. However, is that what the story is all about? Professor Edelman pointed out the risk associated with blind acceptance of the AI technology as headlined in countless media articles. Blind faith in the AI technology might not only be detrimental to social order but also interfere with the development of the AI industry that could bring the humankind more actual benefits.

He quoted the famous science fiction writer Arthur C. Clarke: Any sufficiently advanced technology is indistinguishable from magic. Indeed, every time the late Steve Jobs presented new products, the public and media raved over them, hailing those products to be "magical." How Korea media and the Korean public responded to AlphaGo defeating Lee Sedol who had dominated the Go world for more than a decade in 2016 was not an exception.[6] However, such an attitude leaves a big blind spot behind. For example, not only the public but also even experts have not reached any consensus or

5) For details on the disruptive innovation concept, see Clayton M. Christensen · Michael E. Raynor · Rory McDonald, "What Is Disruptive Innovation?", Harvard Business Review (2015. 12.).

6) AlphaGo has three versions: AlphaGo Lee that defeated Lee Sedol in March 2016; AlphaGo Master that defeated Ke Jie who was the world's no. 1 Go player in March 2017; and AlphaGo Zero that was recognized as the world's strongest Go AI by winning 89:11 against AlphaGo Master solely based on self-learning. At the end of that year, the Alpha Zero came on the scene, which mastered not only Go but also many other games based on a single self-learning algorithm. For details, see David Silver et al., "Mastering the Game of Go with Deep Neural Networks and Tree Search", Nature 529 (2016); David Silver et al., "Mastering the game of Go without human knowledge", Nature 550 (2017); David Silver et al., "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm", arXiv:1712.01815 (2017).

suchlike on the question: “What is AI?”[7)]

Notwithstanding, we are using the AI technology, whether knowingly or unknowingly, and live by the benefits it brings us. Edelman mentioned examples of the use of the AI technology: enhanced spam mail filtering, improved security using biometrics, more accurate photo tagging, and more accurate predictions based on inter-relations of behavioral patterns. In the past AI was primarily used for Internet ads, but now its applications cover a wide range of areas. For instance, there is not much difference in what schools look like between 1919 and 2019. However, AI-based pedagogy has enabled novel, tailored learning methods that would allow us to acquire a large amount of knowledge that otherwise would take us a long period of time. Hospitals have harnessed the power of AI to significantly curb the occurrence of sepsis as a fatal side effect that costs many lives. Another well-known example of AI application is to use AI to optimize school bus allocations and service routes.

However, he highlighted that today’s AI does not have a free will or a sense of identity as seen in the Terminator, nor is it almighty, and we are still in the early stages of technical development. Whereas the computer has developed for more than 50 years, the new AI methodology represented by deep learning (DL) was established only about five years ago.[8)] Still, headlines bombing on us every day might lead us to a misunderstanding that AI has and acts with human-like autonomy. However, in an AI system many human intentions are incorporated, including that of the designer. Accordingly, humans, rather than AI itself, should be held accountable for any errors and discriminative behav-

7) For this reason, in their world-famous AI textbook, Russell and Norvig mentioned four AI approaches - thinking humanly, acting humanly, thinking rationally and acting rationally - rather than trying to define what AI is. See Russell and Norvig, *Artificial Intelligence: A Modern Approach* (3rd Edition), Prentice Hall (2009).

8) Of course some may see the history of AI as long as that of the computer, or other may think that the two have the same origin (initiated by Alan Turing and pioneers). However, the AI in Edelman’s context should be seen in the narrow sense, that is, the DL technology that made a meteoric rise with the emergence of the AlphaGo. Then the standpoint to consider AI outrunning humans for the first time at the Large Scale Visual Recognition Challenge by ImageNet in 2015 to be a milestone in the development of AI would have a point.

iors by AI arising from wrongful learning.[9)]

Along with the accountability issue, another malaise humanity may face, when we see the AI technology as a magic, is the problem of agency. As well known, AI is a very complicated being even technology experts are unable to have a thorough knowledge of it. Hence it is often likened to the "black-box."[10)] Of course, one may find other objects in our daily lives such as automobiles, mobile phones nothing less complicated than AI. But, many are familiar with these technologies and at least have an abstract understanding on how they work. However, (DL) AI tells us few things other than phenomenal facts that certain results were derived from big data analysis. We do not need to understand how everything surrounding us in our daily lives works, of course, and doing so does not necessarily mean we are genuine "agents" in our lives. However, blindly accepting what significantly affects rights and freedoms of humanity as something magical without an attempt to understand the concrete operation is nothing but the loss of agency.[11)]

Edelman mentioned AI-based decision making on credit rating as a typical example. On the one hand, AI enables credit rating based on non-financial information for those who have no previous credit history, for example utility payment history, telephone bill payment history, etc. On the other hand, the characteristic of AI decision making based on correlations rather than causality would mean that it may take factors that humans would find irrelevant, for example if someone charges his/her mobile phone regularly, as a key variable for its decision making.[12)] Even though such a decision might be more

9) Concerning errors in AI, see Anh Nguyen · Jason Yosinski · Jeff Clune, "Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images", arXiv:1412.1897v4 (2015). For discrimination by AI, see Solon Barocas · Andrew D. Selbst, "Big Data's Disparate Impact", California Law Review Vol. 104 (2016).

10) This is so-called opacity in AI. For a well-known precedent study, see Jenna Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms", Big Data & Society (2016).

11) For example, choice architecture or echo chamber shows how individual choices are influenced by external factors in their daily lives and politics, judgment thereon should depend on whether such decisions are based on informed consent. See Michal S. Gal, "Algorithmic Challenges to Autonomous Choice", Michigan Technology Law Review Vol. 25, Issue 1 (2018), p. 63.

12) For the limitations of DL AI that depends not on causality but on correlations, see Gary Marcus, "Deep

accurate than decisions made by humans, its justification is another issue that requires separate considerations.

Of course the conventional legislation has not sat on its hands. For example, the U.S. Fair Credit Reporting Act provided the right to correct credit rating based on incorrect information and the non-discrimination obligations. It did not specifically envisage the emergence of the AI technology, but issues can be resolved by rational interpretation of the legal principles such as fairness and minor amendments in response to anticipated vacuum in regulations. Edelman stressed that what is important here is "governance" that goes beyond AI ethics, finding the difference between ethics and governance comparable with the difference between political philosophy and regulation. As discussions in political philosophy are ultimately embodied in the form of regulations, AI ethics, too, will need to be embodied in the form of governance at some point. For example, whether the right to explanation can be derived from the European General Data Protection Regulation (GDPR), or whether to strengthen explanatory clarities in exchange for declines in accuracy even in areas where accuracy would be more important than explanations requires a social consensus.

In addition, Edelman highlighted the necessity for collaboration, as legal systems are likely to converge into one point given the characteristics of the AI technology despite differences in culture and environments among countries. Such discussions over legal systems are of extensive value, as they are not a zero-sum game in essence. This justifies many governments, NGOs, academic circles, and businesses concentrating their efforts on discussions over AI governance for years on the contrary to the common idea considering AI to be the part of competition, especially "arms race."[13] He concluded his speech by highlighting that in order to make that happen communications between technology and policy experts are essential.

Learning: A Critical Appraisal", arXiv:1801.00631v1 (2018), pp. 12-13.

13) As a notable example, figures involved in many different fields related to AI gathered in Asilomar, where discussions over bioethics were initiated in 1975, in January 2017 and announced the 23 AI principles. For more details on these ethical principles, see Future of Life Institute, "Asilomar AI Principles" (2017).

2. Keynote Speech II

– Procurement as Policy: Administrative Process for Machine Learning

The second keynote speech by Professor Mulligan was centered around governance where the government opts for AI (here the focus was down to machine learning), in other words, public contracts and procurement.[14] As we know, public decision making finds root in the principle of democracy. Hence, it should reflect social values and be kept in check by the rule of law and other social restraints to ensure the right decision. Meanwhile, the government's decision making should be visible and underpinned by rationality and expertise, and ultimately by public engagement and supervision. In particular, it should be contestable by citizens when they become subject to the unfair execution of state authority. The right of access to courts and due process are not fully materialized until then, which in turn leads to social trust.

Mulligan mentioned the Eric L. Loomis case, which triggered disputes surrounding this issue. Loomis was sentenced six years in prison at the first trial with reference to the case management and decision support tool called COMPAS (Correctional Offender Management Profiling for Alternative Sanctions). He could not have access to information on how the system reached the conclusion, the inputs as the ground for the conclusion, or any weights applied in the system. The government did not have such information, and the maker of the COMPAS system was protected by the intellectual property laws, hence not obliged to offer such information. He claimed that the sentence was against due process and unlawful for many reasons. But the Supreme Court of Wisconsin rejected his argument.[15] Mulligan pointed out that at the core of this case was the "opacity" in AI-based decision making, which could spill over into all conducts of the state using AI.

---

14) Deirdre K. Mulligan · Kenneth A. Bamberger, "Procurement as Policy: Administrative Process For Machine Learning", Berkeley Technology Law Journal Vol. 34 (2019).

15) State of Wisconsin v. Eric L. Loomis, 2016 WI 68, 881 N. W. 2d 749 (2016).

A problem concerning this is that AI algorithms are not necessarily neutral or objective. Many situations, including defining relevant variables and weights and deciding on whether to select or employ certain models, may pose policy problems. For example, in the credit rating issue mentioned by Professor Edelman, the variable "charging the mobile phone" premises at least economic power to purchase a mobile phone, hence not considered value neutral. Likewise, we have found no straightforward answers to questions, for example, to what extent we should trust outcomes resulting from applying data, which contain humans' free will and other social characteristics, to algorithms (that are primarily for predictions in the field of natural science) and how we should interpret them.[16] For instance, if setting thresholds for face recognition, sensitivity levels should significantly vary when applied to general purposes and when used for identifying criminals. In this sense, AI-based decision making involves some kinds of valuation, like whether to provide outputs from the algorithm to the users in the as-is state or to offer opportunities to revise them.

Mulligan pointed out that AI is misunderstood to be value neutral as it is subject to procurement, so its politics- and policy-related aspects remain not fully magnified.[17] Such an issue further stands out in the context of the topic of this speech, that is, government procurement. Many people see the procurement of AI not more or less than the procurement of any other private goods. But as mentioned earlier, AI algorithms should not be considered as simple commodities as they serve as a medium for us to understand and form the world. In this sense, in the AI procurement process careful considerations should be given to the rightful incorporation of values such as transparency, fairness, and democracy, in addition to general evaluation factors such as price and

16) The predictive security software PredPol and other AI algorithms are based on the similarities in the occurrence patterns between crime and earthquake. See George O. Mohler et al., "Self-exciting point process modeling of crime", Journal of the American Statistical Association Vol. 106, Issue. 493 (2011).

17) AI programs and AI-embedded hardware are economic products and at the same time products of science and technology. In this sense, AI, too, should be brought to a series of philosophical questions, for example the objectivity and value neutrality of science and technology, the confrontation between technological determinism and social determinism surrounding the relationship between science and technology and social development. For details, see Maarten Franssen · Gert-Jan Lokhorst · Ibo van de Poel, "Philosophy of Technology", Stanford Encyclopedia of Philosophy (2018).

performance. However, working-level officers responsible for public procurement may well lack understanding on the AI technology and its political implications. This implies the viewpoints of private enterprises are incorporated in AI policies, so that administrative services are at the disposal of the private judgment of technological experts that lack democratic legitimacy. It could pose a risk for the infringement of procedural rights of individuals involved.[18)]

In principle, the courses of administrative conducts that affect the rights and interests of citizens are subject to a range of procedural protections. Such procedural provisions ensure, to some extent, the legitimacy and legality of administrative dispositions and allow for ex post problem-posing, thereby strengthening accountability. Mulligan asserted that if AI's decisions have significant impact on individuals' rights and interests, rather than simply being used as a tool, these strict requirements should be satisfied. Furthermore, she took notes on discretion granted to the administration and argued that administrative agencies are allowed to and should make adjustments in various system standards.

Achieving democracy in AI-powered administration requires the involvement of various expert groups to cope with the aforementioned issues. Mulligan invoked the U.S. National Taxpayer Advocate system, which provides a feedback to AI used by the Internal Revenue Service. Finally, Mulligan concluded that employing "contestable AI" designs that allow for retort by those facing problematic situations through impact assessment, prototyping, and simulations would present a way to achieve the transparency of AI procurement and administrative procedures accompanied.

3. Panel Discussion

The panel members for Session 1 included the session's keynote speakers Professors Edelman and Mulligan, Professor Hyunseop Kim (Seoul National University Department

18) As a pioneering study that coined "technological due process", see Danielle Keats Citron, "Technological Due Process", Washington University Law Review Vol. 85 (2008).

of Philosophy), Jake Lucchi from Google, Ministry of Science and ICT Director General Kyung Hee Song, and Jeongwon Yoon from Amazon Web Services. The debate started with Professor Kim's argument that, in terms of "trust" as the keyword in the first session, the roles emotions play in the human's cognitive process may serve as a clue. One of the factors that allowed the Deep Blue to defeat the chess champion Garry Kasparov was the fact that the Deep Blue was not affected by the swirls of emotions as we humans are. In this sense, some questions that unemotional machines might be able to bring outstanding outcomes even in moral decisions beyond humans' imagination.[19] Professor Kim, in response, acknowledged that emotions sometimes prevent us from making the right decision, while pointing out that emotions can also serve as a mechanism for us to incorporate information disregarded by conceptional evaluations in our decision making.

In philosophy, emotions are not confined to the subject's sensations or feelings towards the object but include elements of evaluation. From the former viewpoint, the evaluating of them would be unfeasible, but from the latter viewpoint, one could distinguish "right" emotions from the "wrong." For example, suppose a person felt the sense of fear by looking at an object. Whether his feeling of fear was right or wrong would depend on if the object was a real snake or just a toy. In this context, the emotion-driven cognitive process may be a valuable mechanism to acquire information complementing non-emotional cognitive processes. If one feels sudden-attacked by fear while he roams around an area described to be safe on the Internet, such a feeling might be an unsubstantial error or it might be indicative of a fact based on which he could save his life. This leads us to an idea that we need to appropriately incorporate cognitive systems based on intuitions or emotions in automated decision making processes, rather than trying to completely exclude them.

Still, Professor Kim was clear that he tried to highlight that emotions would give us

19) For example, see Colin Allen · Gary Varner · Jason Zinser, "Prolegomena to any future artificial moral agent", Journal of Experimental and Theoretical Artificial Intelligence Vol. 12, No. 3 (2000), p. 260.

momentum to ponder on and carefully consider things, rather than intending to indicate emotional superiority over reason.[20] He admitted that AI's decision making systems are not necessarily separate from humans' emotional systems at all times, and the emotional systems may be incorporated in AI's decision making process by conceptualizing its roles.[21] Then Kim concluded his speech by presenting the following considerations: initially AI-human collaboration was highlighted as means to strengthen humans' agency or AI's accountability. However, one may find a more substantial reason to do so, that is, to harness emotions to improve AI's capacity for automated decision making.

Then Jake Lucchi mentioned Google's governance principles to improve public trust.[22] Last year Google established a series of guiding principles on what to achieve with the AI technology and what to avoid[23], and among them Lucchi put emphasis on the company's efforts to overcome bias. For example, Google's "What-If" tool is intended to of-

---

20) Such a viewpoint is in line with Kahneman's dichotomy perspective to divide the modes of thinking into the fast, instinctive and straightforward "System 1" and slower, deliberative and complicated "System 2" and emphasize the roles of the latter. For more details, see Daniel Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux (2011), Part 1.

21) It also has implications for discussions over the recent topic surrounding the Artificial Moral Agent (AMA). As a notable study on the AMA, see Wendell Wallach and Colin Allen, *Moral Machines: Teaching robots right from wrong*, Oxford University Press (2008).

22) Google, "Perspectives on Issues in AI Governance" (2019. 1. 22.).

23) Sundar Pichai, "AI at Google: our principles" (2018. 6. 7.).
Here Google highlighted social benefits, avoiding bias, safety, accountability, privacy, and scientific excellence.

1. Be socially beneficial.
2. Avoid creating or reinforcing unfair bias.
3. Be built and tested for safety.
4. Be accountable to people.
5. Incorporate privacy design principles.
6. Uphold high standards of scientific excellence.
7. Be made available for uses that accord with these principles.

And Google found don'ts in the areas of risk and harm, weapons, surveillance, and contravention of international law and human rights.

1. Technologies that cause or are likely to cause overall harm. Where there is a material risk of harm, we will proceed only where we believe that the benefits substantially outweigh the risks, and will incorporate appropriate safety constraints.
2. Weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people.
3. Technologies that gather or use information for surveillance violating internationally accepted norms.
4. Technologies whose purpose contravenes widely accepted principles of international law and human rights.

fer “counter-factual explanations” by providing intuitive illustrations on what results a model would yield when a certain data point changes.[24] In addition, the tech company also employs a range of tools to see if certain data sets are over- or under-represented, or apply pre-defined policy restrictions to models. Improving transparency or clarity, as mentioned by Edelman, tolls accuracy to some extent. However, Lucchi found that, considering other important principles held by Google, for example accountability, and if taking a balanced viewpoint, the profit-seeking business could find decreases in accuracy, to some extent, justifiable.

Of course, balancing between many different principles is required, and overly highlighting certain values should be avoided. For instance, as well noted, improvements in transparency lead to increased vulnerability to gaming by malignant actors.[25] Considerations should be given to trade-off between different principles, for example a test taken with the aim to improve fairness might infringe an individual’s privacy. In addition to the abovementioned technical tools, Google is making multi-faceted endeavors to take a right view to many elements to consider. They include supportive training, exclusive staffs and governance processes like efforts to ensure incidents promptly reported to the superiors and efforts to ensure these ideas to be incorporated in corporate culture. Finally, Lucchi emphasized that, to ensure effectiveness applying these efforts to the real world, it would be important to give considerations to concrete contexts on the use of AI and pursue harmonization with national and international norms.[26]

Kyung Hee Song from the Ministry of Science and ICT briefly mentioned what the Korean government is doing in the course of building public trust in AI. One of the im-

24) For details on the “What-If” tool, see http://pair-code.github.io/what-if-tool/. Counter-factual explanations tell us what variables we need to change, and how much, if we want to produce a desired result on the premise of the status quo. As a notable study on counter-factual explanations, see Sandra Wachter · Brent Mittelstadt · Chris Russell, “Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR”, Harvard Journal of Law & Technology (2018).

25) Jenna Burrell, op. cit., pp. 3-4.

26) Lucchi particularly emphasized the case of Singapore. For details, see Singapore Personal Data Protection Commission, “A Proposed Model Artificial Intelligence Governance Framework” (2019. 1.).

portant features of AI is increased uncertainty. As noted, there are rosy prospects about the future of AI, accompanied by the gloomy future scenarios, for example discrimination, unemployment, etc. The public and private endeavors in international society and ethical principles mentioned above can be understood as part of attempts to minimize such adverse effects. It is not to mention that this could not be done by a single player or area. Rather, it is a problem of humanity that should be tackled by the concerted efforts of all stakeholders. The OECD's ethical guidelines and the G20 Summit Declaration are some of notable products.[27)]

To keep up with such endeavors by international society, the Korean government has developed ethical guidelines while focusing on AI R&D. In line with the international discussions, the guidelines are represented by the so-called "PACT" principles: Publicness, Accountability, Controllability, and Transparency, which are applicable to the trihedral agents, i.e., developers, suppliers, and users, respectively. In addition, the national AI strategy is forthcoming: currently the government is collecting opinions from various stakeholders. Also, it is developing a plan to invest KRW 2,500 billion in the coming five years with the aim to build an ecosystem for AI R&D. This would include multi-faceted efforts such as opening a graduate school of AI studies, offering project-oriented education, and building hubs. Finally, the government is pursuing more robust international cooperation by seeking international exchanges with France, Russia, UK, Brazil, and more.

Lastly, Jeongwon Yoon presented the standpoint of Amazon Web Services (AWS) to the use of AI. To achieve the goal of improved customer satisfaction, individuals' rights should be protected, and abiding by the positive law in each country should be a prerequisite to technological innovation. To do so, AWS has a policy that limits the use of the company's products (AWS Acceptable Use Policy) and procedures to report suspected violations and abuses in place.[28)] He also mentioned that AWS is seeking constant com-

27) For details, see OECD, "Recommendation of the Council on Artificial Intelligence" (2019. 5. 22.); G20, "G20 Ministerial Statement on Trade and Digital Economy" (2019. 6.).

munications with interested parties by taking part in various open communities. In particular, AWS has developed concrete guidelines on face recognition, which poses significant risk for the breach of fundamental rights, prescribing when not to use it and when to limit its use.[29]

He also pointed out the need to fractionate the characteristics of user groups with the aim to achieve AI democracy in the true sense. The explainability in its true sense is to provide groups of experts, groups with a certain level of understanding, for example data scientists, and groups of the public with different tools. According to Yoon, AWS has made a series of efforts, which allowed the company to customize 97% of technologies tailored to customer demands.[30] As a notable example, 10% of cancer patients in the USA benefit from tumor diagnosis programs powered by AI, as shown in the study by Fred Hutchinson Cancer Research Center.[31] Such an experience has significant implications for the importance of considering customer needs in future AI legislations.

II. Session 2 – Fairness in AI: What Does It Mean, and How Can It be Implemented?

1. Keynote Speech III

– Fairness, Explainability and Trustworthy AI: Technical Challenges and State of the Art

While the first session was themed around "trust," the second session focused on "fairness." How can we secure fairness as an important premise to improve trust in AI?

---

28) For details on the AWS policy, see: https://aws.amazon.com/aup/. For the AWS service abuse report process, see: http://pages.awscloud.com/rekognition-abuse.html.

29) Michael Punke, "Some Thoughts on Facial Recognition Legislation" (2019. 2. 7.).

30) As presented at Amazon's annual conference "re:invent." For details on re:invent, see: http://reinvent.awsevents.com/.

31) For more details, see Taha A. Kass-Hout · Matt Wood, "Introducing medical language processing with Amazon Comprehend Medical" (2018. 11. 27.).

In his keynote speech, Professor Fredrik Heintz presented his answers to this question based on his experience in the European Commission High-Level Expert Group on AI and the group's deliverable, the "Ethics Guidelines for Trustworthy AI."[32)]

According to Professor Heintz, we are at an important point of inflexion when it comes to the proposition of "trustworthy AI." Once the public reach a conclusion that AI is not trustworthy, it would be beyond revoke. After all, the trustworthiness issue is overarching for both businesses and governments wishing to take advantage of the positive sides of AI. Against this backdrop, the European Union is seeking, by fully mobilizing many different methodologies developed by humanity to date, to reach an AI worthy of public trust and beneficial for the human race. The EU guidelines state that trustworthy AI should be lawful, ethical, and technologically and socially robust.[33)] These three elements are mutually complementary, as, for example, AI might bring unintended harm even if it meets the lawfulness and ethics requirements.

Based on the fundamental rights as grounds for trustworthy AI, the guidelines present four ethical principles - respect for human autonomy, prevention of harm, fairness, and explainability - from which seven key requirements are derived: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being; and accountability. Furthermore, the guidelines include an assessment list to ensure AI systems meet the above, consisting of 130 questions. These arrangements are only provisional and will be subject to subsequent complementation, for example ongoing pilot projects.[34)]

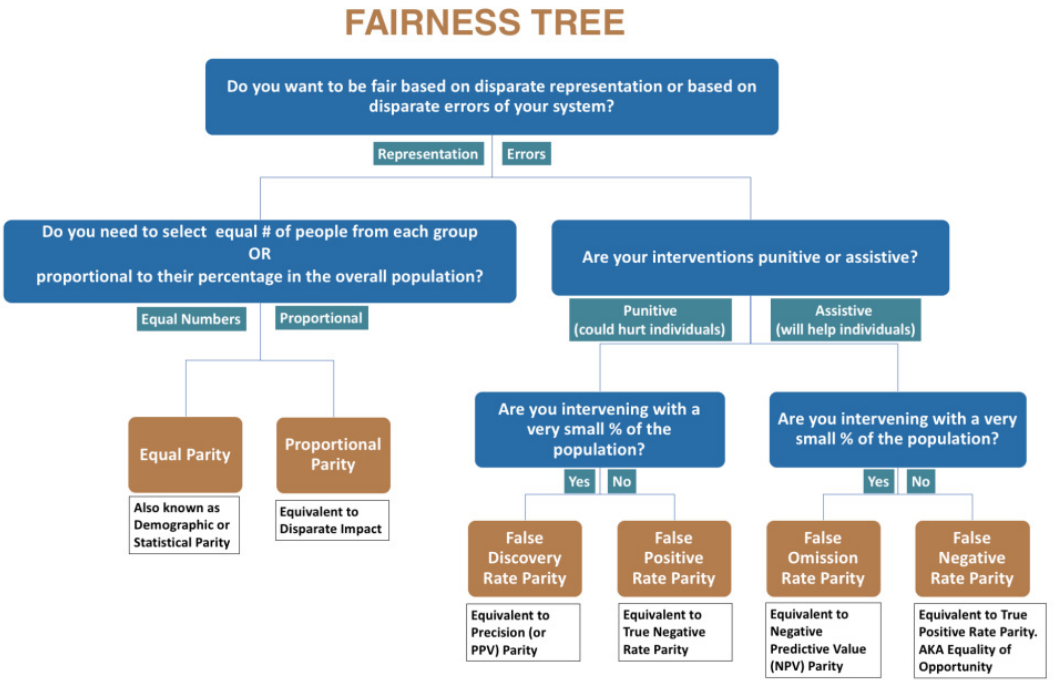
After giving an overview of the report, Heintz furthered discussions over fairness and explainability. To improve fairness, it is essential to think of bias as a hinderance factor

32) High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", European Commission (2019. 4. 8.).
33) High-Level Expert Group on Artificial Intelligence, op. cit., p. 5. This report is more geared towards ethics and robustness than lawfulness.
34) For details, see High-Level Expert Group on Artificial Intelligence, op. cit., p. 9 ff.

against it. As noted in the previous session, bias caused by data sets failing to correctly represent the population often distorts the reality. However, such a problem can largely be resolved by increasing the number of samples collected. In this sense, it could be a practical challenge but one may not see it to be a fundamental problem. On the contrary, situations where bias in the population itself, for example data that reflect historic discrimination in the past, comes into question are much more challenge to resolve. It may break no squares if the ultimate objective of AI is to simply analyze the reality, but it may not be the case if fairness matters and the reality cannot be accepted in the as-is state. Still, assessing or revising outcomes from AI for fairness' sake may pose new problems concerning appropriateness, i.e., who has the authority to do so, and to what extent.

Furthermore, different people have different views as to what fairness means. Even if they can reach a consensus, the question on how to incorporate humans' intuitive understanding in AI as a product of engineering still remains. For example, the Center for Data Science and Public Policy at the University of Chicago offers an open source bias audit toolkit called "Aequitas", and the fairness tree used in this toolkit provides for six types of fairness.[35] Either of those six is not always right or wrong, which essentially calls for valuation depending on specific contexts. Once we have defined what fairness is, we can adjust our decision making process as follows: to put it simple, one or more elements in the input and output processes are revised. For example, pre-processing by partially modifying data in terms of availability or content, or making ex ante or ex post interventions in the algorithm.

35) For Aequitas, see: http://www.datasciencepublicpolicy.org/projects/aequitas/.



[Figure 1] Fairness Tree

(Source : University of Chicago Center for Data Science and Public Policy)

Then Heintz continued with discussions over explainability. Though being an interesting topic on its own, explainability also is a way to improve fairness. He said, as noted in the previous session, there is a trade-off relationship between explainability and accuracy, but it does not seem that the future will be dominated by AI that is only accurate while lacking explainability. Accuracy can be associated with specific contexts in which AI is used, and it will be difficult for AI to earn trust if it, being extremely accurate in some situations and not being so in others, lacks explainability. It does not necessarily lead us to the conclusion that in all situations the use of AI should be backed up by considerable explainability. He pointed out that it would not be the case even for human decision making.

As mentioned by Mulligan, the most important role of explainability is to ensure contestability against unsatisfactory decisions, and "explanations" in this context should be understandable to humans. In this sense, disclosing to the public, who are unable to spe-

cifically understand how AI works, the entire source codes may result in improved transparency, but it does not mean improved explainability. This justifies the Explainable AI (XAI) Project by the U.S. Defense Advanced Research Projects Agency (DARPA) seeking to combine artificial neural networks with highly explainable models, employ highly explainable methods to generate models in the first place, or strengthen visual user interfaces to ensure intuitive understanding.[36] What all those endeavors ultimately aim at is to ensure explainability while keeping accuracy to the greatest possible extent.[37] Improved explainability in AI also helps developers improve systems. Lastly, Heintz concluded by highlighting the need for a transition from correlation-based approaches to causality-based approaches and mentioning that such studies would also contribute to improving our understanding on what fairness is.

2. Keynote Speech IV

– A Responsible Development of AI: With an Example of Federated Learning on Privacy

The last keynote speaker Blaise Agüera y Arcas touched on how to resolve issues surrounding fairness and privacy in AI from an engineer's perspective. To justify the rise of the fairness issue, Agüera y Arcas gave an overview of the difference between how the post World War II computer technology and the recently emerging DL-based AI work. Written in computer program languages, (source) codes basically take sequential operations following strictly defined rules.[38] In that course, errors may occur for many reasons, and the longer the codes are, the harder to read and debug. By performing iter-

36) The DARPA's explainable AI project was already discussed at the second conference. For the more information, see David Gunning, "Explainable artificial intelligence (XAI)", Defense Advanced Research Projects Agency (DARPA) (2017).

37) As a notable model, Heintz mentioned the LIME (Local Interpretable Model-agnostic Explanations, LIME), which aims to grasp the relations between independent and dependent variables by giving some perturbations to the inputs and reading local changes shown in the outputs. For more details, see Marco Tulio Ribeiro · Sameer Singh · Carlos Guestrin, " "Why Should I Trust You?": Explaining the Predictions of Any Classifier", arXiv:1602.04938v3 (2016).

38) It is a distinctive feature of the Turing machine. As a classic paper by the deviser of the Turing machine. For more details, see A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", Proceedings of the London mathematical society Series 2 Vol. 42 (1936).

ative computations fast, computers enrich the life of humanity not only in mathematical calculations but also in many other applications.

In the course of the development of the computer, we could understand what comparative advantages it would have for humans, and what comparative advantages we as humans would have for it. For example, the computer is superior to humans in terms of precision and speed for mathematical calculations, as seen in the cases of the Deep Blue or the AlphaGo. Not to mention it is far less vulnerable to physical limitations such as sleepiness, hunger, or emotional turmoil. On the other hand, humans have superior abilities in terms of fundamental motor functions, i.e., to walk and run.[39] Furthermore, unlike the computer, humans are better at analogy and flexible thinking, meaning they can understand objects or concepts they have never seen before and think creatively. That was the reason why it was challenging to realize Turing's ideal of the "thinking machine" with conventional hard coding.[40]

Developed in recognition of such a problem, the artificial neural network (ANN) represents a methodology in contrast with the conventional hard coding method. The ANN basically mimics the neutral network structure in the brain, i.e., the neuron-synapse connections.[41] Today's deep neural network models, commonly called deep learning, were developed by elaborating the early ANN models such as single-layer perceptron. Recapitulating the human brain, the DL model works superb in pattern recognition the old computer models found formidable, and people are increasingly believing that further developing DL might end up leading to a "thinking machine" like humans do. Agüera y

39) It is commonly known as Moravec's paradox, named after robot engineer Hans Moravec who first indicated that problem. One possible explanation is that, unlike complicated cognitive functions such as advanced reasoning, simple motor functions were included in the course of human evolution long ago. Although AI excelled the human visual recognition at the Large Scale Visual Recognition Challenge by ImageNet, humans still outperform AI in terms of the total sum of energy consumed for visual recognition.

40) A. M. Turing, "Computing Machinery and Intelligence", Mind, New Series, Vol. 59, No. 236 (1950), p. 433.

41) The 1943 model by McCulloch and Pitts is known as the initial idea about the ANN, but the perceptron model coined by Frank Rosenblatt in 1958 is recognized as the first model to encompass learning process. See F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", Psychological Review Vol. 65, No. 6 (1958).

Arcas cited the famous quote by former U.S. Supreme Court Justice Potter Stewart about the controversies surrounding the French movie Les Amants, "I know it when I see it,"[42] to make a point that how DL AI works is becoming increasingly similar to the human intuition. However, as mentioned by Heintz, problems surrounding sample data representation or historical discrimination in population data might lead us to bias. Of course, it can be filtered out, at least partially, by separating training data from validation data or test data, which is an approach actually taken in practice.

Basically, fairness is closely associated with the appropriate implementation of such test processes. Improved representation of training data, separation of training data from test data, and incorporating constraint conditions on inappropriateness in the algorithm are prerequisites for us to make it in line with the ideal of fairness. For example, one may consider adding restrictions on the purpose of the use of AI to the algorithm to prevent sensitive information from being inferred through the facial recognition technology. AI's decision making is capable of not only representing the past but also reproducing past problems in a path-dependent manner, which calls for the need to fundamentally blocking such a problem from occurring.[43] DL requires a huge amount of data and human and physical infrastructure for computation, meaning it develops on the premise of externality affecting the members of the community. Agüera y Arcas emphasized that we need to humble ourselves and admit the fundamental problem that we are unable to reach a perfect social consensus on fairness but take a progressive approach to resolving the most problematic cases in reality.[44]

As technical solutions to privacy issues in the use of DL AI, Agüera y Arcas mentioned the "Edge TPU" and "federated learning." The Edge TPU is a microchip that is

42) Jacobellis v. Ohio, 378 U.S. 184 (1964), p. 197.

43) Values that matter to humans are incorporated by design or by default. As a notable example, refer to Article 25 of the GDPR (data protection by design and by default).

44) This is so-called "Arrow's Impossibility Theorem", which is mathematical proof that a clear order of preferences cannot be determined while adhering to mandatory principles of fair voting procedures and non-dictatorship. See Kenneth J. Arrow, "A Difficulty in the Concept of Social Welfare", *Journal of Political Economy* Vol. 58, No. 4 (1950).

much smaller than a penny. It allows for learning and data analysis on personal IoT devices, rather than Google and large enterprises' massive data centers. Currently data collected by sensors attached to personal devices are uniformly transmitted to central data centers and the results of learning are fed back. Not only is it an inefficient structure that consumes excessive electricity and network traffic, it also involves higher risk for hacking. Federated learning represents a local AI methodology developed as a solution. When federated learning is deployed, primary learning is done on individual devices, and data from the learned model, rather than training data, are encrypted and sent to the data center, which then merges the model data and learn.[45] Agüera y Arcas concluded his speech by highlighting that the Edge TPU and federated learning represent promising alternatives to improve the protection of personal information while pursuing the technological development of DL.

3. Panel Discussion

The panel members for Session 2 included the keynote speakers Heintz and Agüera y Arcas, Norberto Andrade from Facebook, Gary Chan from Singapore Management University, Malavika Jayaram from Digital Asia Hub Hong Kong, and Indrė Žliobaitė from the University of Helsinki. First, Andrade mentioned that Facebook, with the aim to realize fairness in AI, takes a holistic approach to three key elements: the workforce, data, and algorithm. The workforce consists of three teams, which are code developers, data learners, and reviewers. The latter include a legal team, and their roles include coordination with academia and public and private organizations and incorporating their views in engineers' production. Doing so allows for sort of compliance, as in the "privacy by design" mentioned by Agüera y Arcas. All matters considered in the process are documented. Here, considerations are given not only to compliance with legal requirements but also many different types of public demands. As human bias may permeate into the data and/or algorithm, the sensitivity of those involved plays an important role

45) For the Google Edge TPU, see http://cloud.google.com/edge-tpu/?hl=ko. For Google's federated learning, see http://federated.withgoogle.com/.

in improving fairness.

For data, the primary task should be to clarify the objectives of the model and sort out data accordingly. One should be able to report problems in data and identify from what sources the data in question were introduced. This should be accompanied by labelling and taking note of the types of data pre-identified as problematic, as well as separate de-identification measures for privacy protection. In reality, however, it is unfeasible to get rid of all biases in the data level, which should be taken care of in the algorithm level. For example, Facebook unveiled the Fairness Flow tool at its annual technical conference "F8." It classifies subgroups based on the features of data sets and offers at-a-glance visual insights on difference in results and accuracy between them.[46] Still, Andrade emphasized that technical tools cannot be the cure to all social problems and, as mentioned earlier, processes should play important roles. He concluded by highlighting the need to establish best practices on fairness, document successful problem-solving cases, and pursue constant cooperation with external communities.

The next speaker, Chen developed discussions with a particular focus on employment. We all do job hunting, hire someone, or will end up doing any of those activities. In this sense, employment is a universal topic, and AI-related issues have significant impact on the life of humans. The use of AI brings advantages to the employment process that was exclusively taken by humans, such as immediacy, efficiency achieved by assigning human workforce to other duties, and elimination of human bias and many other vulnerabilities. In this context, AI can be used in a wide range of applications including job classifications, goodness-of-fit tests in document screenings and interviews, and career coaching for employees. However, as noted by other speakers, AI's decision making is not perfect and may be prone to many types of bias and intentional and unintentional discrimination. A widely known example is gender discrimination by AI in the selection of job advertisement targeting.[47] To make corrections, the Constitution, statutory laws,

46) For more details, see Jerome Pesenti, "AI at F8 2018: Open frameworks and responsible development" (2018. 5. 2.).

and ethical principles envisaged in guidelines presented by, e.g., the Institute of Electrical and Electronic Engineers or relevant communities may serve as references.

As concrete measures for correction, Chen presented the following corresponding to the abovementioned data and algorithm levels in the context of employment. First he pointed out that job skills should be described in inclusive languages, for example universal skills necessary regardless of gender rather than job skills a specific gender is known to be more skillful at. Achieving representation of training data by subgroup or getting rid of pre-identified negative features is a solution applicable to areas other than employment. Algorithm-wise, ex post audit by trusted third parties (TTPs) and the addition of "randomness" are presented as possible methodologies. In addition, there is a need to provide those who suffer from disadvantages in AI-powered employment processes with intuitive explanations, for example counterfactual explanations mentioned by Lucchi. Furthermore, it is important to keep them updated, rather than stopping them as one-off measures.

The next speaker, Jayaram, raised a fundamental question about the notion of "bias." In general, bias is associated with stereotypes and prejudice in personal perception, hence understanding in the social and systemic levels possibly being disregarded. In this sense, more social terms should be used, for example racism, or structural oppression.[48] In this context, Jayaram pointed out the need for multi-disciplinary studies on AI beyond the boundary of technology and presented a study that represents successful collaborations.

The gist of the study mentioned by Jayaram is that fairness is a concept that forms part of a wider system beyond the domain of technology, hence approaching the realization of fairness only from technological perspectives would not lead us to the right solution. The authors of the paper categorized errors in the course of technological ab-

47) Ads containing information on high-pay jobs were significantly more exposed to males compared to females. For more details, see Amit Datta · Michael Carl Tschantz · Anupam Datta, "Automated Experiments on Ad Privacy Settings", Proceedings on Privacy Enhancing Technologies Vol. 2015, Issue. 1 (2015).

48) See Kinjal Dave, "Systemic Algorithmic Harms", Data & Society (2019. 5. 31.).

straction of the concept of fairness into five types: failure to model the entire system over which fairness will be enforced caused by missing important human and physical elements (framing trap); failure to understand AI that works fair in a certain context may not work well when applied to a different context (portability trap); failure to account for the full meaning of the social concept of fairness through mathematical formalism (formalism trap); failure to understand how the insertion of technology changes the behaviors and embedded values of the pre-existing system (ripple effect trap); and failure to recognize the possibility that the best solution to a problem may not involve technology (solutionism trap).[49] Jayaram concluded her speech by pointing out that a range of issues, including technology gaps between advanced and developing countries, and the impact of AI-human coexistence on child upbringing and development, should be included in the discussions over fairness.[50]

The last speaker, Žliobaitė, found it encouraging that studies on fairness in AI have surged since the last year and mentioned some points to consider. First, she took the example of the internationally controversial Boeing 737 MAX crash. The lesson we should take from this tragedy is that when we design a technological system, it involves certain presumptions about human behavior to use it. A system with inappropriate presumptions may pose higher risk to be exposed to unexpected incidents. She pointed out that AI may act more consistently than humans but being consistent does not necessarily mean being objective.

That leads us to the need to interpret outcomes AI offers. A problem here is that, as mentioned earlier, humans and AI have different advantages in different areas, and there are inherent limits in ex post validation by humans in areas where AI excels humans. Furthermore, we need to understand the difference between causality-based human deci-

49) Abstraction commonly used in computer science inherently tends to remove much information associated with social contexts, while such information is often considered pivotal to achieve fairness. See Andrew D. Selbst et al., "Fairness and Abstraction in Sociotechnical Systems", Proceedings of the Conference on Fairness, Accountability, and Transparency, ACM (2019).

50) For the criticism mentioned by Jayaram, see Berkman Klein Center, "IDRC Global Symposium on AI & Inclusion Outputs" (2018).

sion making and correlation-based AI decision making even in areas under control by humans' ex post validation and acknowledge the limits of AI decision making that is unavoidably exposed to bias. That would explain the reason why the term "raw data" is an oxymoron.[51] Finally, Žliobaitė concluded by highlighting the importance of multi-disciplinary approaches, for example expressing social values in technological constraint conditions.

51) See Lisa Gitelman (eds.), *"Raw Data" Is an Oxymoron*, The MIT Press (2013).

Preparation & Support Team

Co-Directors of the SNU AI Policy Initiative : Haksoo Ko and Yong Lim

Preparation Team Leader : Taehoon Kim

Preparation Coordinator : Jonggu Jeong, Ji Hoon Park, Hyeln Kim, Eunsoo Kim

Documentation and Report Compilation : Do Hyun Park

Interpretation and Guide Services for Foreign Participants:

Seong Goo Kang, Yujin Kwon, Na Hyeong Kim, Dong Yeon Kim, Se Young Kim, Sion Kim, Jin Woo Kim, Junseong Ma, Yujun Park, Jin Sik Shin, Heesung Yoon, Hyunseo Lim, and Wonhwi Cho (Seoul National University Law School AI and Law Society)

Video Production : Ji Hyun Lee, Hyunuk Kang, and Youngchae Cho (Seoul National University Department of Undeclared Majors), and Hojae Byun (Seoul National University School of Dentistry)

Simultaneous Interpretation : Ji Eun Chun and Sunhee Sohn