

## Beijing Forum 2019 Humanity and Legal Governance in the Digital Era (III)

On Sunday morning, the last meeting of “Humanity and Legal Governance in the Digital Era”, Panel Session 12 of the Beijing Forum (2019), was held in Meeting Room No. 3, Yingjie Exchange Center, Peking University. There were two sessions, and five scholars were invited to give presentations.

The topic of the first session was “Ethical Principles in Global Governance of AI”, and was moderated by professor Huw Price, from the University of Cambridge. Scott J. Shackelford and Rachel D. Dockery together gave a presentation, titled “Governing AI: Insights from Polycentric Governance”. They first quickly reviewed the concept and development of AI nowadays, the benefits of AI and meeting the challenges from AI. Then they made a careful survey of current AI governance efforts and turned toward a polycentric model based on it. Taking different AI governances on autonomous vehicles as a case study, they showed how to apply such social science constructs to the world of AI. Considering polycentric governance, they provided some implications for policymakers, most important of which was that cybersecurity should be taken as social responsibility, and governments should actively participate in the UN’s efforts to promote global cybersecurity and cyberpeace and to avoid the use of cyberspace for conflict.

Professor Zeng Yi, from Peking University, then spoke on the topic “AI Ethical Principles: Global Landscape and Technical Challenges”. He began with a brief reflection on whether current so-called AI is really AI. He introduced ten major AI principles to provide a global vision of AI development. Although it seems in general that these principles cover everything important, after a careful check it was found that no principles actually covered all those fields concerned. This is a critical issue, and Prof. Zeng suggested we need to concentrate in some more fundamental way for a better understanding. Therefore, he proposed a “Human-AI harmony society”. He took the case of Beijing AI Principles and stressed that the main idea is really not about which country is going to be the first, but on a shared future together. How do we do that? By studying the case of emotion recognition in classrooms, which shows the possibility of a disharmonious human-AI relationship. Zeng wanted to say, rather than just have AI do something according to moral principles, we need

to have AI learn models for ethics. Is AI with the self too risky? Completely to the contrary, as he suggested: it is AI that can learn that may contribute to the harmonious relations we wish for and believe in for our future. After the report, Thorsten Jelinek, from the Taihe Institute, questioned Zeng on whether he has presupposed the concept of our self-consciousness before we figure out what it is. Professor Zeng responded that the case in which robots learn doesn't mean they necessarily have consciousness. Rather, we are actually challenging the cognitive science on what consciousness should be.

The next session, "Transformation of Ethics", was moderated by associate professor Liu Zhe, from Peking University. He first referred to Mark P. Mckenna earlier presentation on "Technology and the limits of ethical reasoning". It is universally presupposed in AI discussions that new technology will raise new ethical problems. However, Liu thinks that is actually not true. New technology, like AI, as he suggests, never raises new problems. The challenges that arise from them can actually be explained by the proposition that new technology pushes the boundary between human and technology and sharpens existing ethical problems. Humans are concerned with the meaning of being and of being a human, and, exactly as suggested in Professor Zeng's response to Jelinek above, development of AI makes this problem much more significant. He took data used in advertising as an example to argue for his idea mainly through the perspective of legal discussions. Believing that new technology doesn't shape us in different ways, he sincerely invited us to concentrate on the fundamental values of being human.

Professor Mark Coeckelbergh, from University of Vienna, then spoke on the topic of "Ethics of AI: A Discussion of the Problem of Responsibility". He first argued that there is no such thing called general AI. What needs to be paid attention to is the existing and to-exist problems of AI around us. Besides, he agreed with Professor Mckenna that none of the ethical problems of AI are newly born. Rather, they are old problems that revive again under the development of new technology, such as the problems of privacy and security. He went on to discuss the problems concerning responsibility related to AI, and expressed his view that AI shouldn't be responsible itself but rather it should be those who build the AI who should be responsible, for AI doesn't have control condition and knowledge condition that functions as responsibility conditions that an agency should have, according to Aristotle. He then also provided some ideas on how to deal with the ethical problems of AI.

# BEIJING FORUM

## 北京论坛学术简报 ACADEMIC BULLETIN OF BEIJING FORUM

Finally, professor Huw Price, from the University of Cambridge, gave a concluding report for the session, titled “The Future of Artificial Intelligence: Academia’s Role in Getting in Right”, in which invited the audience to get better involved in this topic. He first suggested that AI is a long-term revolution, and has been influencing our life and will continue to influence it in our future. Next, he analyzed four major reasons explaining why the correct development of AI should rely on academia: (a) only universities have the breath of expertise needed; (b) few other organizations have the luxury of long-time horizons and some independence from the commercial and government world; (c) academia is global and (d) academia is a resilient source of epistemic authority. Further, he put forward six design principles for better academia, which are: (a) epistemic humility; (b) interdisciplinarity; (c) not just academia; (d) connectedness across borders and boundaries; (e) disciplinary identity and (f) the long view.