



AIIG
 清华大学人工智能国际治理研究院
 INSTITUTE FOR AI INTERNATIONAL GOVERNANCE
 TSINGHUA UNIVERSITY



AIIG
 清华大学人工智能国际治理研究院
 INSTITUTE FOR AI INTERNATIONAL GOVERNANCE
 TSINGHUA UNIVERSITY

Outcome Document International AI Cooperation and Governance Forum 2021

January 2022

Outcome Document
International AI Cooperation and Governance Forum 2021

January 2022

Content

I. About the forum	3
II. Forum Outcomes	4
III. Summary of Remarks	11
(1) Opening remarks	11
(2) Report-Main Plenary I	19
(3) Report-Main Plenary II	25
(4) Report-Main Plenary III	32
(5) Reports-Thematic Sessions	41
IV. Next Step	71

I. About the forum

The 2021 International AI Cooperation and Governance Forum, hosted by Tsinghua University in support of the United Nations Development Programme, was held from December 4-5, 2021. In this year's forum, leading AI experts and practitioners from around the world focus on the theme of “how to build a balanced and inclusive AI system,” and thereby promote the construction of a shared future for mankind.

The two-day forum is being held in a hybrid online and offline format. More than 70 experts from over 20 countries attended this two days conference across three plenaries, seven sub-forums and the U&AI Bootcamp special event while the number of on- and offline participants exceeded 500. More than 160 news reports covered the event and the live broadcast of our forum reached 3 million views.

II. Forum Outcomes

(1) A basic consensus on international cooperation on AI governance is reached at the forum.

First, it is essential to promote win-win cooperation on AI governance and in other fields. AI governance has become an important global issue, and the international community must join hands to address the risks and challenges brought on by the development of AI. Communication among countries on AI regimes and policies needs to be enhanced, and forward-looking dialogue aimed at addressing new challenges needs to be actively pursued.

Second, a formal, standing, and multistakeholder international dialogue platform on AI governance needs to be established for regular discussion, joint research, and multistakeholder adoption. The United Nations has played a very important role in reaching agreements and consensus among stakeholders, and the international dialogue platform on AI governance should facilitate cooperation—across borders, generations, sectors, and

disciplines—in the UN framework to develop an appropriate sub-framework for global AI governance.

Third, cross-cultural trust in AI ethics and governance needs to be built. The first step should be to facilitate global collaboration on the cultural and conceptual aspects. The next step is to understand, through cross-cultural cooperation, the use and regulation of AI in different cultures, so that our perceptions are not confined by geopolitical boundaries. Finally, a cultural-ethical balance must be struck between universality and differentiation in promoting cross-cultural, international cooperation on AI governance based on mutual respect, enhanced trust, and inclusive dialogue.

Fourth, it is necessary to promote the establishment of consensual AI ethical standards and governance guidelines. This process requires consultation, mutual enrichment, cross-cultural trust, respect, inclusiveness, as well as a comprehensive approach to development and security and to efficiency and equity, so as to avoid inadequacies, cherry-picking, discrimination, and exclusivity in the policy orientation of the governance guidelines and to build a balanced and consensual global framework of AI governance

guidelines and ethical standards.

Fifth, bilateral and multilateral cooperation mechanisms need to be actively explored. It is important to fully ensure the potential and participation of third-party countries other than China and the United States in order for the international community to jointly build an inclusive AI governance model.

(2) The forum proposes to develop a balanced and inclusive AI governance system.

First, the forum lays out a basic framework of an AI governance system. In terms of values, it should be built on the preparedness for the worst-case scenarios, and be human-centered and development-oriented, with a global vision. In terms of governing bodies, the framework should engage multiple stakeholders including governments, businesses, civil society, universities and research institutions, media, NGOs, and international organizations. The stakeholders should have different authority, resources, interests, and restrictions in the socio-technical system of AI, and should interact extensively through formal and informal channels

to form a governance assemblage. There are five targets of AI governance. The first is data, or more specifically, the security of data itself, as well as the autonomy, controllability, and macro security of data. The second is algorithms, or more specifically, their stability, security, explainability, and equitability. The third is computing power. In this regard, multilateral win-win cooperation shall be sought to avoid technology monopoly. The fourth is external environments, which concerns ways to compensate or remedy the externalities of AI technologies, especially the negative ones, as necessary. The fifth is scenarios, referring to the need to consider a variety of specific scenarios of AI applications.

Second, the forum proposes AI governance mechanisms of multiple paths and levels. The first is to reach a basic value consensus. The most central element of the distinctive AI principles set out worldwide is steering AI development on the value basis of inclusiveness, sharing, prudence, and responsibility. The second is to promote collaboration among stakeholders while they work in their respective spheres of competence. Broadly speaking, there are four types of governing entities: governments, AI technology providers, AI technology users, and the public and third-party

organizations. For the governments, they need to empower AI technology providers (including R&D organizations, platform organizations, universities, etc.) while maintaining effective oversight. For AI technology providers, they need to proactively and effectively interact with and empower the governments so that the quality and efficiency of public services are improved. For AI technology users, they ought to cooperate and interact with technology providers, while accepting government regulation. The public, third-party organizations and other civil society entities need to actively provide timely, effective, and comprehensive scrutiny over the governance elements, targets, and processes. Thirdly, the governance mechanisms require ongoing iterative optimization. Many early technologies can only reach 70% or 80% of social penetration after decades of diffusion; but the rapid innovative development and application of AI means that this technology moves faster than its governance guidelines. This calls for a combination of governance mechanisms and models: one is exploratory or forward-looking governance, which means that we set up protocols for calculated risks; the other is adaptive or agile governance, which effectively balances the quality and pace of governance.

Third, the forum proposes a diverse governance toolkit. A multidimensional portfolio of governance tools at the “micro, meso, and macro” levels shall be adopted in developing a comprehensive AI governance framework to promote the healthy development of the technology.

(3) The forum explores into the future of governance in the metaverse.

First, the metaverse should be human-centered. The virtual world should be developed to help and empower the real world, not to replace it.

Second, the governance of the metaverse should be forward-looking. The transcendent conception and construction of the metaverse means that if information and data continue to serve a key path to the making of the world, the metaverse will lead to a new type of Internet applications and a new shape of society where new technologies are integrated and where reality and virtuality merge. In the future, the target of metaverse regulation and governance may not be the Metaverse Empire—a single-

dimensional technological system—but a metaverse republic or a consortium of microcosms—a multidimensional, integrated technological system. There is a need to revisit the impact of the currently dispersed but connected virtual spaces on the real world.

Third, the metaverse will require shared governance. At present, the overall development of the metaverse is in its infancy, with a low level of technological application. An approach to metaverse governance where research, development, and governance are conducted simultaneously can be adopted to ensure that no significant risks arise. In the future, metaverse governance should be built on sufficient research and extensive, multistakeholder discussion to achieve a balance between development and governance and avoid a one-size-fits-all governance framework.

III. Summary of Remarks

(1) Opening remarks

Wang Yongqing, member of the 19th Central Committee of the Communist Party of China and vice chairman and leading Party member of the 13th National Committee of the Chinese People's Political Consultative Conference, states that China's AI governance has and will always put people at the center, aim to deliver for humankind, encourage technological innovation, adopt a balanced approach to development and security, and promote fairness and justice. Aligned with the AI governance principles of the United Nations and the G20, China's AI governance aims to contribute to the betterment of humankind and to sustainable economic, social, and ecological development. **The Chinese government advocates, as always, advancing international dialogue and cooperation on AI development and governance, addressing global digital governance issues through joint consultation, development and sharing, and working together toward an open, fair, just, and non-discriminatory environment for digital development. In the face of the risks**

and challenges posed by AI development, the international community must join hands to address them and improve global AI governance. We need to promote openness and inclusiveness, uphold the vision of building a community with a shared future for humankind, reject the logic of hegemony and ideological bias, maximize the openness and inclusiveness of cooperation mechanisms, visions and policies, and build a responsible regime of global AI governance. We also need to promote consultation and mutual enrichment, promote universal values including peace, development, fairness, justice, democracy, and freedom, build cross-cultural trust, show respect and tolerance for differences, and adopt a balanced approach to development and security as well as to efficiency and equity, so as to avoid inadequacies, cherry-picking, discrimination, and exclusivity in the policy orientation of the governance guidelines and to build a balanced and consensual global framework of AI governance guidelines and ethical standards. We need to promote shared governance and create open platforms for all stakeholders to engage in AI governance. Through dialogue and cooperation—across disciplines, sectors, regions, and borders—we ought to forge multistakeholder

compliance with the fundamental ethics of science and technology so that progressive values are embedded into the mathematical models of AI and AI is used for good; in this way, the best practices of AI governance can be promoted, and shared AI governance with active multistakeholder participation will be achieved.

Qiu Yong, president of Tsinghua University and academician of the Chinese Academy of Sciences, notes that AI is the landmark technology of the Fourth Industrial Revolution, and intelligence has brought far-reaching changes in how we work and live over the recent years. The era of intelligence is here. However, while AI is improving social productivity and modernizing governance, it also brings new challenges, such as public security concerns, privacy infringement, the digital divide, the accountability conundrum, ethical failures and so on. AI governance requires multistakeholder collaboration, multi-dimensional objectives, integration of pluralistic values, and strengthened global cooperation and dialogue on AI governance, so as to build a balanced and inclusive AI governance system. As important incubators of scientific and technological progress, universities need not only to make

breakthroughs and innovations in AI theories and technologies, but to advance technology for good, and shape good AI values and ethics, in order to promote the establishment of a scientifically sound international mechanism of AI governance.

Ms. Beate Trankmann, UNDP resident representative for China, points out that AI is being leveraged in the fight against the pandemic and climate change. At the same time, human biases can potentially find their way into AI algorithms and incentivize infringement on individual privacy. **As the application of AI becomes more and more omnipresent across the world, international cooperation will be essential to build consensus on common approaches to AI governance that are ethical and fair. Therefore, it will be critical to ensure that the global conversation on AI is open to all countries.**

Li Meng, vice minister of science and technology of China, states that China is committed to fostering technology for good and attaches great importance to AI governance. In recent years, the Ministry of Science and Technology has released several guidelines on AI governance, including the *Governance Principles for the*

New Generation Artificial Intelligence and the Ethical Norms for the New Generation Artificial Intelligence. In rolling out these principles, the ministry set up a mechanism for coordinated promotion, carried out social experiments, engaged businesses and local entities, and actively participated in international governance cooperation. **He stresses that future AI governance should be agile so as to balance AI development with safety and security; it should be adaptive and inclusive so as to ensure balanced benefits sharing among stakeholders; it should be sustainable so as to achieve intelligent and green development; and it should advance the vision of building a community with a shared future for humankind so as to strike a balance between development and win-win cooperation.**

Xu Xiaolan, vice minister of industry and information technology of China, says that the combinations of the Internet and AI and of the industrial Internet and AI has transformed the way we live and produce respectively. AI will also have a profound impact on the modernization of social governance. China attaches high priority to the development of the AI industry, and the Ministry of Industry and Information Technology (IIT) has ensured the implementation

of the policies and deployments of the State Council and the CPC Central Committee by identifying key areas to make AI breakthroughs, setting up national manufacturing innovation centers for intelligent sensors, intelligent connected vehicles and other sectors, and promoting the establishment of pilot zones of AI innovation and application in Beijing and seven other cities. In the future, **the IIT will continue to cooperate with relevant parties on the sustainable and sound development of AI, and on the creation of inclusive AI governance norms, striving to open up a new landscape of international cooperation. The IIT is ready to contribute more Chinese wisdom to the establishment of international AI governance institutions and generate more practical results for international scientific and technological exchanges and cooperation.**

Ma Shengkun, deputy director-general of the Department of Arms Control of the Ministry of Foreign Affairs of China, states that China attaches great importance to the security of AI governance, and that the Chinese government has all along coordinated and advanced the development and risk control of AI technology with a responsible attitude. China has also issued ethical principles and

norms for researchers, and clarified supervision mechanisms and requirements in a timely manner. In the future, measures should be taken to regulate AI-empowered military applications. **AI should abide by international law and fundamental norms governing international relations. AI ethics should follow the vision of AI for good. The evaluation and supervision of AI technology should be strengthened. And in terms of governance, global engagement must be advocated for.**

Maria-Francesca Spatolisano, officer-in-charge of the UN Office of Secretary-General's Envoy on Technology, notes that AI is one of the key digital challenges identified in the UN Secretary-General's *Roadmap for Digital Cooperation*. As a technology that can bring about meaningful changes, AI comes with daunting challenges. **Better global guidance and cooperation are needed to maximize the benefits of AI.** This point was reinforced in the UN Secretary-General's report *Our Common Agenda*, where he stressed AI as an area where greater global regulation might be needed and proposed establishing a new multistakeholder body on global AI cooperation, with a particular emphasis on engaging the Global South. **The world could and must work together to create inclusive,**

responsive, and effective global cooperation structures that can meet the challenges posed by AI.

Andrew Chi-Chih Yao, Turing Award winner, academician of the Chinese Academy of Sciences, foreign member of the National Academy of Sciences, dean of the Institute for Interdisciplinary Information Sciences at Tsinghua University, and chair of the Academic Committee of the Institute for AI International Governance (I-AIIG) at Tsinghua University, points out that as AI becomes embedded in our economy and society, it is increasingly important to maintain a positive interaction and synergy between technology and socio-economic development, which requires a balance between the development, application, and governance of AI. However, AI has also posed challenges to privacy protection and data security in the era of digital economy.

Ever-refining algorithms, efficient computing power, and privacy protection have charted new courses for research on AI theories. The establishment of a balanced and inclusive AI governance system will unlock greater potential and ensure stronger security of AI.

(2) Report-Main Plenary I

Main Plenary I is themed “To Build a Balanced and Inclusive AI Governance System.”

Xue Lan, distinguished professor of Arts, Humanities and Social Sciences at Tsinghua University, dean of Schwarzman College, dean of I-AIIG, and chair of the National Expert Committee on AI Governance of China, states that **AI governance has progressed from a pure concept to the forefront of on-the-ground practices**, and is facing challenges in terms of data, algorithms, computing power, and scenarios. There will be a need for comprehensive AI governance systems that are built through multistakeholder approaches that include governments, businesses, civil society, universities and research institutions, media, NGOs, and international organizations. In terms of values, the systems should be built on the **preparedness for the worst-case scenarios, and be human-centered and development-oriented, with a global vision**. Specifically, the systems will exercise governance over algorithms, data, computing power, external environments, and scenarios using a multidimensional governance toolkit at the micro,

meso, and macro levels. Ultimately, the systems are expected to **promote the healthy development of AI and deliver for humankind through governance mechanisms such as “establishment of value consensus, multistakeholder collaboration and synergy, and iterative optimization of governance visions,”** among others.

Gong Ke, president of the World Federation of Engineering Organizations, executive director of the Chinese Institute of New Generation Artificial Intelligence Development Strategies, and academic member of I-AIIG, believes that an AI governance system should, first and foremost, be based on safeguarding human rights and achieving sustainable development. Second, it should aim to promote rather than curb the innovative development of AI. Third, it should be an open symbiosis of multistakeholder governance. Fourth, it should adopt a variety of tools including technologies, regulations, and education. Fifth, it should be effectively implanted into AI platforms. **As an emerging productive force, the development of AI is unstoppable. Therefore, AI governance should aim to advance AI innovation for the good of humanity and our planet, not to contain it by**

confining it to the existing systems.

Haoliang Xu, UN assistant secretary-general and UNDP director of the Bureau for Policy and Programme Support, notes that **AI is already furthering the Sustainable Development Goals across the world and has become a driving force in other sectors as well, from education to transportation. Our governance of this technology needs to be developed in concert with AI.** Just as AI can help to scale up programs, it can scale up bias and exclusion too—through lack of knowledge, negligence, or malicious intent. **We need more cooperation—across borders, sectors, and generations—to develop the necessary governance frameworks,** such as multistakeholder approaches that include governments, international organizations, tech companies, and civil society. Adopting an inclusive approach can build trust, flag risks in areas such as discrimination, and speed up implementation.

Wendell Wallach, Carnegie-Uehiro senior fellow at the Carnegie Council for Ethics in International Affairs and academic member of I-AIIG, believes that we need a new 21st-century approach to international cooperation on AI governance. An international AI

governance mechanism should be agile, adaptive, anticipatory, responsive, and inclusive. Chinese and American participation in developing international governance schemes for AI will contribute to greater feasibility of such schemes. **Countries need to manage the inherent tension between the need for national security and the need for international cooperation, accepting international standards while actively engaging in forward-looking dialogue aimed at addressing new challenges.** AI governance standards are being developed and implemented globally in large numbers, and standards can clash between countries. This requires respect for cultural differences and effective international AI governance mechanisms as the technology evolves.

Su Jun, director of the Think Tank Research Center of Tsinghua University, dean of the Institute of Intelligence Society Governance of Tsinghua University, and deputy director of the Management Committee of I-AIIG, believes that AI and other fast-evolving emerging technologies are creating far-reaching impact to the economy, social governance, and people's lives around the world. An intelligent society is facing technological, market, social, and

cognitive risks, which calls for AI governance systems with human values. This requires scientific approaches to technology-induced challenges. **As we return to experimentalist governance—a traditional public management philosophy, we follow a scientific, normative, quantitative, and evidence-based research paradigm, taking advantage of the extensive and diverse AI applications in China in the social experiments on AI, so as to explore the Chinese path of intelligent society governance.**

Rohinton Medhora, president of the Center for International Governance Innovation, highlights the importance of governance over the data input-technology output process. In terms of data in AI governance, we need to establish globally accepted ethical standards as the value foundation of AI, universal standards on data use to ensure data security, as well as an algorithmic accountability mechanism to ensure that data is put to good use. At the technical level, we need a more efficient global intellectual property system to improve data sharing and intellectual property mechanisms, and a global database to drive innovation with new paradigms. **In the future, it is important to set up international cooperation**

mechanisms for AI that include governments, the private sector, and NGOs, so as to fully ensure the potential and participation of third-party countries other than China and the United States. In addition, a people-centered global AI working committee should be established to promote the right values.

On “promoting international AI cooperation and governance,” the speakers agreed that an international AI dialogue platform should be established as soon as possible. The platform should be designed in the frameworks of the United Nations and the existing dialogue mechanisms, and bring together experts across sectors and disciplines to promote comprehensive research on AI governance.

(3) Report-Main Plenary II

Main Plenary II is themed “Frontier Technologies in Artificial Intelligence and Governance.”

Susan Athey, member of the National Academy of Sciences, member of the American Academy of Arts and Sciences, winner of the John Bates Clark Medal, economics of technology professor at the Stanford Graduate School of Business, and associate director of the Stanford Institute for Human-Centered AI (HAI), believes that **universities have a critical role to play in steering AI innovation. As it is now difficult to develop explainable algorithmic frameworks due to a lack of ethical and philosophical training for technical practitioners in the private sector, furthering relevant research could play an important role in identifying problems, developing frameworks for practices, and guiding AI governance.** In addition, the current platform economy characterized by “software-as-a-service” has been widely embraced thanks to the scale effect of data. The demand for AI and data may create a “pseudo” market concentration, and the prediction of a future where “machines

replace human” is highly challenging, prompting us to rethink how AI may be used to address public management issues such as population aging and to make AI-based public services more efficient.

Yolanda Gill, former president of the Association for Advancement of Artificial Intelligence (AAAI) and academic member of I-AIIG, believes that the lack of human knowledge about intelligent mechanisms, the complexity of intelligent behavior itself, the limited means of observation, and the differences in individual knowledge, professions, beliefs, and cultural backgrounds have led to an array of challenges in AI research, highlighting the need to strengthen fundamental AI research, which requires efforts across sectors and disciplines. **At present, understanding AI mechanisms and building AI world models are the two major challenges facing AI research. On the one hand, understanding AI mechanisms requires the development of a “perceive-decide-act” intelligence model and deeper understanding of the mechanisms of the mind. The practices of neuroscience research consortia might prove to be a useful model for the establishment of a global AI database and research community.**

On the other hand, an AI world model needs to be built on human experience, social customs, and expertise, so it would be ideal to create a free collaborative knowledge base to promote knowledge sharing at the global level through the participation of all.

Zhang Bo, academician of the Chinese Academy of Sciences, honor director of the Institute for Artificial Intelligence at Tsinghua University, and academic member of I-AIIG, points out that the unexplainability of deep learning and other algorithms has led to fairness and security concerns as well as unreliability and untrustworthiness in the first two generations of AI. The key to developing third generation AI is to develop explainable and robust AI theories and methods, to develop safe, trustworthy, reliable, and scalable AI technologies, to build deep learning platforms that support explainable algorithms through data- and knowledge-driven approaches, and to empower AI with optimized security and defense. Siphoning true intelligence from data depends on the support and guidance of knowledge, and requires effective regulation on the use of data, so as to promote the innovative development of AI by integrating knowledge, data, algorithms, and

computing power.

Gao Wen, academician of the Chinese Academy of Engineering, dean of the School of Electronics Engineering and Computer Science of Peking University, director of Peng Cheng Lab, and academic member of I-AIIG, believes that AI development has reached a critical phase—from new generation AI to strong AI—and by 2030, China is expected to become a global leader in overall AI development. **From a strategic perspective, China, the US, and Europe are in fierce competition regarding AI talents, research, development, application, hardware, and data. Today, China’s strengths in AI lie in its strategic policies, data resources, application scenarios, and potential talents, while it is relatively weak in basic theories, original algorithms, key components, international platforms, and number of senior talents. From a tactical perspective, AI 2.0 needs to adopt statistical AI based on big data to meet the needs of large-scale AI applications, and to encourage exploration into the territory of strong AI. Possible technical roadmaps for this include “explainable machine learning + reasoning” and “bionic systems + AI high computing power.” From a security perspective, the major**

risks of strong AI stem from the unexplainability of models, unreliability of algorithms and hardware, and uncontrollability of autonomous consciousness. Therefore, AI 2.0 should apply DPI and “leakproof” technologies to ensure data security and privacy protection. Furthermore, it is crucial to explore solutions to AI ethics, and adopt different risk prevention strategies at different stages in the “theory-technology research-application” cycle.

Harry Shum, foreign member of the National Academy of Engineering, foreign fellow of the Royal Academy of Engineering, adjunct professor of Tsinghua University, and former executive vice president of Microsoft Corporation, notes that AI is already at work all around us, even in legal practice. For example, many courts in the US are using machine learning and AI for criminal sentencing, including deciding important issues such as lengths of sentences. However, AI is already making decisions we don't understand. Just seeing the “black box” of AI decision-making is not enough; we need to look inside and understand the specific contents and causalities in order to achieve explainable AI. At the same time, responsible AI should ensure fairness, reliability,

privacy, inclusiveness, transparency, and accountability. For an emerging industry such as AI to deliver for humanity, it needs to draw on the lessons learned by other sectors.

Ya-Qin Zhang, foreign academician of the Chinese Academy of Engineering, dean of the Institute for AI Industry Research at Tsinghua University, and academic member of I-AIIG, believes that carbon neutrality represents yet another transformation in the energy structure. It is an inevitable choice toward sustainability, and an opportunity for industrial restructuring and development, pressuring businesses to transform their models and increase efficiency. AIoT—the integration of AI and the Internet of Things—can empower green computing and **help realize net zero in three ways: first, data-driven and AI-optimized engines can achieve intelligent decision-making; second, it enables multi-parameter, whole-process optimization of system configuration; and finally, intelligent perception could be achieved through the integration of multi-source, multi-dimensional and heterogeneous perceptions. AIoT can be used for promoting energy integration, reducing carbon emissions in the ICT industry, and fostering emerging industries.**

The participants agreed that it is crucial to promote win-win cooperation, regular discussion, joint research, and multistakeholder adoption in AI governance and other fields. It is also essential to develop consensual AI ethical standards and governance guidelines, enhance global governance capabilities in algorithms and data security, and work toward an inclusive AI governance model.

(4) Report-Main Plenary III

Main Plenary III is themed “The Future of Governance in the Metaverse.”

Duan Weiwen, director and fellow of the Academy of Philosophy of Science at the Institute of Philosophy and director of the Research Center for Science, Technology and Society, both at the Chinese Academy of Social Sciences, believes that the technological world that human beings are building for themselves is built on cyber-physical systems (CPS), and **whether it is called the metaverse or not, if information and data continue to serve a key path to the making of that world, it will lead to a new type of Internet applications and a new shape of society where new technologies are integrated and where reality and virtuality merge.** An issue that needs urgent and in-depth reflection is that the transcendent conception and construction of the metaverse may lead humanity into a hyperspace which is not only a world of perceptual experiences such as reality, virtuality, and extended reality, but also a world that transcends reality and virtuality, with idea- and meaning-creating functionality such as signs/symbols

and imagination. When considering the higher dimension of signs/symbols and imagination, one can rely more on a combinatorial approach in constructing the metaverse, which means to combine people's autonomous consciousness and conceptual creation with appropriate technologies, without necessarily moving toward accelerationism—relying solely on technology for infinite upgrades—or even falling into a single convergent model. If this is recognized, **the target of future metaverse regulation and governance may not be the Metaverse Empire—a single-dimensional technological system—but a metaverse republic or a consortium of microcosms—a multidimensional, integrated technological system. The governance of the metaverse should be forward-looking.**

Shen Yang, director of the New Media Research Center of the School of Journalism and Communication, Tsinghua University, believes that the metaverse is a new world characterized by the fusion between reality and reality, and that it has realized the expansion of space-time, which entails multiple risks, such as disputes over the ownership and attribution of responsibility of

virtual humans, and individual cognitive and behavioral alienation in human-computer interaction. At present, the overall development of the metaverse is still in its infancy, with a low level of technological application. An approach to metaverse governance where research, development, and governance are conducted simultaneously can be adopted to ensure that no significant risks arise as it advances the growth of the metaverse industry. In the near future, this issue can be discussed within the framework of existing systems, laws and policies, and individual power and interests should be the current priority of metaverse governance. **Future metaverse governance should be built on sufficient research and extensive, multistakeholder discussion to achieve a balance between development and governance and avoid a one-size-fits-all governance framework.** Problems need to be identified, analyzed, and solved as the metaverse evolves to achieve targeted and dynamic governance.

Zeng Yi, co-director of the China-UK Research Centre for AI Ethics and Governance at the Institute of Automation, Chinese Academy of Sciences, and chief scientist at I-AIIG, believes that the current concepts concerning the metaverse are still

scientifically unclear and misleading. It is especially important not to push the younger generation into this unknown when the basic concepts, visions, and applications of the “metaverse” still seem to be at significant risk. At present, the concept is still rather poorly portrayed. Furthermore, the risks triggered by the Internet and AI are already enough to send shockwaves throughout human society even before the arrival of artificial general intelligence, not to mention the risks that will be brought on by the “metaverse,” which is based on the development of modern networks and AI. Finally, from a cultural perspective, both the virtual and real worlds are products of historical development. Virtual societies may be reshaped, but history is indelible. **Cultural enrichment and mutual trust are important “infrastructure” of shared governance on virtual platforms. From a technical perspective, when we promise concepts to the global public and political systems, we must ensure that such fundamental concepts can withstand scientific scrutiny and are technically feasible and sound for societal application. New concepts reinvented in complete isolation from the real world can lead to existential risks.**

According to Wei Qing, CTO of Microsoft (China), everyone has a different understanding of the metaverse. **The metaverse should be human-centered and be developed to empower the real world, not to replace it, and we need to prevent addiction to the virtual world.** It is increasingly recognized that both virtual and real spaces require order and governance. However, there is still no clear answer to the question that requires a universal consensus—what is good governance? The digital divide poses a great challenge as the digital society evolves, and we need to ensure that no one is left behind as technology advances. There needs to be enough confidence that new technologies can and will contribute to the betterment of humanity.

Yu Chen, co-founder and president of YeePay and co-author of *Metaverse Token*, believes that all virtual worlds created can be regarded as metaverse, and virtuality does not have to be something intimidating. The history of humanity itself is moving from reality toward virtuality. The root cause of many problems in the metaverse and modern society is that symbols are often confused with their referent objects and mimesis with reality. Virtual spaces need order and governance. The possible threats of new

technologies cannot be simply generalized and analogized to those of previous technologies; after all, AI is a technology that could potentially destroy humanity even without consciousness—the secondary hazards of new technologies are often underestimated. **Technology is, in principle, neutral, but it is indeed an “amplifier” of human nature, which calls for dynamic governance of new technologies.** The metaverse is creating value possibilities for humankind, and we need to weed out the bad and leave the good.

According to Zeno He, industry business development manager at NVIDIA Omniverse, metaverse is a concept, not a technology, which has been under heated discussion as new digital technologies emerge and are put to extensive use. Specifically, metaverse is the fusion of AI, machine vision, physical simulation, high performance computing, and many other technologies, forming a new industry segment where reality and virtuality merge. **The prefix “meta” in “metaverse” stands for the harmonious coexistence of the objective world and the subjective consciousness and will of humans, which should aim to serve the real world.** One example is robot waiters that provide food

recommendations for customers based on their preferences. Over the past decades, most people have experienced the dividends of the virtual world. But interestingly, if one can distinguish the differences between the real world and the metaverse, they often choose to improve their real life by starting to work out to achieve a balance between the virtual life and the real life.

Zi Min, a Bilibili KOL, believes that in the metaverse, we are more than just a virtual identity, but a collection of many other things. Today, many may not be aware that they are already using the metaverse, or they simply do not understand the concept of metaverse, which is mainly based on virtual and real technologies. The virtual world is just beginning to take shape, while the real world is already established. It is important to keep in mind that **the virtual world is developed to empower the real world**, and that multistakeholder cooperation is crucial for the metaverse to deliver for those in need.

According to AI and crypto artist Song Ting, **the metaverse is not a substitute for the physical universe, nor is digital art a substitute for physical art; new cultures based on digital twins**

hold great promise in the metaverse. Industry should continue to develop programmable and scalable art spaces so that the metaverse can empower human creativity, not reduce it. Surmounting geographical boundaries, the metaverse enables artists to communicate with their peers around the world. In addition, credit is crucial in the metaverse because blockchain technology makes every transaction visible.

The conception of the metaverse allows for the first time a holistic view of technologies such as the Internet, AI, blockchain, and secure computing that have been developed separately over the past two decades and the various types of virtual systems they have created, revealing how human politics, economics, and life have been reorganized by virtual systems into a large number of digital-based and complex relationships that have overturned the established power, rights, and relationships in the real world. The panelists agree that cyberspace has created tensions with the established order, organization, and structure of the real world, and has posed challenges on a whole new dimension to real-world governance. These tensions and challenges have given rise to a series of problems such as challenges to human nature, social rifts,

political unrest, and financial risks, with more issues brewing. **It is therefore necessary to revisit the impact of the currently dispersed but connected virtual spaces on the real world. The issues that need to be urgently addressed for the next generation Internet are the lack of institutional consensus and regimes, as well as the institutional conflicts between virtual and real-world systems.**

(5) Reports-Thematic Sessions

1. Thematic Session #1: Closing Digital Divide in the Age of Global Digital Interdependence

Themed “Closing Digital Divide in the Age of Global Digital Interdependence,” Thematic Session #1 explores how South-South cooperation will help ensure that no country is left behind and help poorer countries improve their AI governance capacity, how global AI governance and digital cooperation mechanisms can be more inclusive, and how fair treatment for the Global South can be ensured through standard setting, knowledge exchange, capacity building, and enhanced digital cooperation.

Eugenio Vargas Garcia, tech diplomat, deputy consul general, and head of science, technology and innovation at the Consulate General of Brazil in San Francisco, USA, and senior adviser and team leader on peace and security, humanitarian and legal affairs at the UN General Assembly, believes that AI should not be given the right to make life-and-death decisions. Developers of AI systems may have only considered this issue from a technical

perspective, not from a socio-cultural one. **Therefore, AI should not be readily applied without considering the broader social context.**

Rohan Samarajiva, chairman of LIRNEasia Think Tanks (Sri Lanka), believes that **dialogue is essential and can even influence policy making to some extent.** It is often the case that the governments make the final decision, excluding all other stakeholders, including civil society and the private sector, from the process. This can lead to policy failures due to knowledge divergences. To address this issue, more extensive consultation is needed.

Vasuki Shastry, associate fellow in the Asia-Pacific program at Chatham House, points out that **explainability and conditionality are crucial.** Explainability means that governments should explain to the general public the scopes, goals, and implications of their AI policies as well as how data is disseminated and collected. Conditionality means that the malicious use of AI should be prevented to protect humans from harm.

Nanjira Sambuli, fellow in the Technology and International Affairs Program at the Carnegie Endowment for International Peace, believes that **the UN is the only international organization that can achieve cross-generational engagement on AI**. For this reason, the UN needs to be ready to lead the conversation.

Nannan Lundin, innovation counsellor at the Swedish Embassy in China, says that **AI governance is a resource-intensive endeavor**, requiring not only the necessary investment of efforts and funding, but also mutual understanding and trust. Meaningful results are possible only after we find out ways to facilitate effective multistakeholder dialogues on AI governance.

The participants believe that AI governance has become an important global issue toward which countries' attention needs to be directed, and that international dialogue and cooperation on AI governance should be actively promoted in the UN framework.

2. Thematic Session #2: Cross-Cultural Cooperation in AI Ethics and Governance

Themed “Cross-Cultural Cooperation in AI Ethics and Governance,” Thematic Session #2 touches on ways to realize sustainable AI through cooperation on AI governance and ethics, the role of culture in such cooperation, and measures to dismantle cross-cultural mistrust and misunderstanding.

Seán S. ÓhÉigearthaigh, co-director of the Centre for the Study of Existential Risk and program director of the Leverhulme Centre for the Future of Intelligence at the University of Cambridge, points out that AI has and will continue to have a significant impact on the world we live in. **Many of the challenges that AI presents to humanity will not vary by culture or country—we all face these challenges, and cooperation is the way out.** However, this will not be easy as people from different backgrounds have different needs and priorities. Cross-cultural cooperation can only be achieved by communicating, understanding, and empathizing with the perspectives, opinions, and concerns of others.

Amandeep Singh Gill, former executive director of the Secretariat of the UN Secretary-General’s High-Level Panel on Digital Cooperation, presents the necessity to **move toward an**

interdisciplinary and inclusive AI collaboration paradigm. One of the challenges we face is the ongoing tension between universal AI and culturally specific AI. The former is aligned with our common humanity. For example, the *Universal Declaration of Human Rights* and other governance-related documents have provided important guidance for AI governance. But AI is also used in specific contexts. Each country, for example, has its own framework, standards, and approach to data collection and governance. There is therefore a need to resolve such contradictions between this universality and locality. Cross-cultural cooperation presents another challenge. Without the enforcement of legally binding rules, it will be more difficult to introduce AI governance treaties. Therefore, it is important to ensure that consensus-based rules can be adapted to each cultural context, and cross-cultural dialogue is needed to steer the process. Third, as AI is culturally distinct in its ability to mimic and understand human intelligence, it is necessary to maintain dynamic agency and cultural sensitivity to understand AI in a culturally sensitive way.

Vincent Müller, professor at Eindhoven University of Technology

and fellow at the Alan Turing Institute, notes that in cross-cultural cooperation on AI ethics, there is a need to analyze, case by case, **where tolerance is called for and where dialogue and criticism are called for.** It is sometimes not easy to delineate and balance these two scenarios, which can lead to the extremes of too much criticism or too much tolerance. **The key to building cultural respect among partners lies in a balance between critical dialogue and tolerance.**

Emma Rutkamp-Bloom, chairperson of the UNESCO Ad-Hoc Expert Group on AI Ethics, says that **it is essential to advance global AI ethics through cultural cooperation, which should be done not through a one-size-fits-all or top-down approach, but through bottom-up collaboration.** AI technology affects all humanity, and its sustainable development requires stronger international cooperation. In addition, **it is important to balance rights-based and duty-based ethics with respect to data ownership, and to adapt practices to different regulatory contexts and values.** It is also important to consistently adopt and support internationally recognized mechanisms, respect internationally agreed mechanisms supported by internationally

recognized values in different contexts, and ensure that all cultures are actively engaged in AI governance mechanisms and lifecycles, so as to integrate cross-cultural cooperation into the governance framework from the bottom up.

Mark Findlay, professor of law and director of the Center for AI and Data Governance at Singapore Management University, points out that there has been more consideration on security and reliability in AI governance and less on ethics, and that AI ethics frameworks are often not seen as regulatory frameworks. **From a cultural perspective, getting the conversation on AI ethics heard by a wider audience requires both culturally appropriate mechanisms and frameworks, as well as conceptual and regulatory frameworks that are universal and global. To achieve this, the first step is to build trust between people and between people and institutions (mechanisms). The second step is to take into account the application of AI in various fields as we work together to develop AI ethical principles.** It should be approached not only from a cross-cultural perspective, but also from the perspective of the global community.

Danit Gal, former AI chair of the Office of the UN Secretary-General and associate fellow at the Leverhulme Centre for the Future of Intelligence at the University of Cambridge, believes that **AI ethics and governance is centered on cross-cultural cooperation.** Despite the growing internationalization of cultures and cross-cultural collaboration, two elements of true cross-cultural cooperation are marginalized: the willingness to communicate and mutual respect. **We need to first promote global cooperation on the concept of culture. Second, we need to advance cross-cultural cooperation to understand the use and regulation of AI in different cultures, so that our perceptions are not confined by geopolitical boundaries. Finally, more time and effort should be invested in building mutual respect and trust to forge cross-cultural cooperation.**

The panelists agree that we humans are inherently interdependent on and interconnected with each other, especially when faced with the global challenges posed by AI. Cultural differences do exist for people living in different cultures, but these divergences are more superficial than fundamental. Therefore, despite the differences in the understanding and perception of AI ethics in different parts of

the world, a universal recognition of fundamental AI ethical principles can still be achieved across the cultural divides, and universal values can be established based on human consensus. To promote the sustainable development of AI, it is imperative to strike a cultural-ethical balance between universality and differentiation in promoting cross-cultural cooperation on AI governance based on mutual respect, enhanced trust, and inclusive dialogue.

3. Thematic Session #3: AI and Climate Action

With the theme of “AI and Climate Action,” Thematic Session #3 focuses on how AI can help address climate change, how to balance AI and climate change and promote relevant innovation and cooperation, and how to promote interdisciplinary solutions to climate change.

Zhang Xiliang, director of the Institute of Energy, Environment and Economy at Tsinghua University, projects **significant changes in future energy management resulting from national policies on renewable energy and climate, among others.** The current

model-based policymaking may well transition into a data-driven approach in the future. This shift will be reflected not only in the power system, but also in future management and policy decisions. Science is critical in how the UN is tackling climate governance issues. **The UN has played a significant role in reaching agreements and consensus among all stakeholders.** Climate change is an issue that presents considerable uncertainty and reflects highly divergent interests among the member states of the UN, and there is also a need to consider the efficiency and cost effectiveness of the climate action. Moreover, countries share the responsibility to reduce emissions. **In this process, special attention should be given to confidentiality issues**—the data in the future global carbon market may have to be open and transparent, but some involves trade secrets and even affects national sovereignty, and it is therefore necessary to find solutions that can provide reliable data while safeguarding trade secrets and national sovereignty. AI and computer science have a lot to offer in this regard.

Wu Libo, associate dean of the Institute for Big Data and director of the Center for Energy Economics and Strategy Studies, both at

Fudan University, argues that **AI application requires computing power, but there are huge regional differences in computing resources.** For example, China's AI-intensive coastal region, where many e-commerce companies are headquartered, has the strongest demand for computing resources, but most of the renewable energy facilities are built in the northwest of the country. As the demand for AI increases along the southeast coast, there is a need to improve the region's energy supply system to make it more effective in combating climate change by optimizing the allocation of computing power between the east and west and integrating the computing resources with the power resources in the east.

According to Felix Creutzig, professor at the Mercator Research Institute on Global Commons and Climate Change, **obligatory data sharing can provide enormous help in creating a good policy environment.** There are still cities without consistent access to important data such as aeronautical and ground transportation maps, even though it is readily available in the databases and useful for social development.

Sylvain Duranton, senior partner and managing director of Boston Consulting Group and global leader of BCG Gamma, believes that after COP26, reducing carbon emissions has become a priority concern for businesses, governments, and individuals. In addition, the growing application of AI will have a strong impact on the environment, and whether the impact will be good or bad largely depends on two things: **whether the evolving AI will grow into a big carbon emitter, and whether it will have a positive or negative impact on the planet.**

Tian Feng, dean of SenseTime Intelligent Industry Research Institute, points out that in order to promote the sustainable development of the industrial metaverse, we should continue to reduce energy consumption and improve energy efficiency in the AI industry and digital economy, and facilitate the green transformation of digital industrialization to develop “green AI” and “green metaverse.” **AI ethics must be people-centered and environment-oriented,** and its minimum targets should include the promotion of sustainable industrial development and controllable sci-tech development.

The panelists agree that AI and other 21st-century frontier technologies will serve as effective tools and innovative paths to achieve the Sustainable Development Goals. AI, as part of a powerful solution, will play a positive role in addressing climate change and other global challenges, but only if it delivers for everyone equally and all countries have a say in its use and development.

4. Thematic Session #4: AI and International Security

Themed “AI and International Security,” Thematic Session #4 looks into two topics: understanding the impact of AI on global strategic stability and exploring ways to alleviate AI-related security risks.

In Session 1, the panelists focus on **understanding the impact of AI on global strategic stability**, sharing insights into how AI affects the strategic stability of major countries around the world in the context of changing major-power relations, and how this would influence the powers that are less competitive in the area.

Directing attention to **exploring ways to alleviate AI-related security risks**, Session 2 facilitates a discussion on taking confidence-building measures to prevent and control the risks associated with AI development and deployment, and on applying AI in international security in a proper and positive manner.

The discussants agree that the application of AI will create considerable uncertainty to the realm of global strategy, and that the international community should continue to improve AI governance and ethics and actively explore cooperative paths to prevent and mitigate AI-related security risks.

5. Thematic Session #5: AI Algorithm and Social Justice

Entitled “AI Algorithm and Social Justice,” Thematic Session #5 brings together computer scientists and sociologists to share their insights into the design, development, and application of AI products, opportunities for interdisciplinary collaboration among scientists, ways to address AI-induced biases, and greater transparency in AI.

In her speech, Chen Ling, director of the Center for Industrial Development and Environmental Governance at Tsinghua University, points out that AI cooperation and governance is a global issue. At present, it is widely believed that algorithms are merely technical tools, and we need to reach a consensus and develop common rules on the use of such tools. It is undeniable that AI algorithms have induced fairness concerns in many fields. However, improving algorithmic fairness necessarily compromises technological efficiency—there is a trade-off between the two.

Algorithmic discrimination stems from three things: data, algorithms, and human behavior. It is essentially a concentration of technological order that cannot be eradicated. Nevertheless, government intervention and remedies are still essential. Furthermore, it is important to recognize the inherent, irreconcilable conflicts among different perspectives of what fairness means: it is impossible to ensure fairness from the beginning all the way through the end. So, how do we choose between conflicting ideas and calculations? The answer is to find acceptable fairness. The starting point for global algorithm governance is to search for a consensus-based, minimum acceptable fairness in algorithms.

Rayid Ghani, director of the Center for Data Science and Public Policy in the Department of Machine Learning at the CMU Heinz School of Information Systems and Public Policy, emphasizes on the importance of defining privacy. **Whether privacy and trust issues will arise depends on how data is used; algorithms do not need data—they are a means, not an end.** If the ultimate issue is about what values an algorithm is expected to convey, there needs to be greater transparency about the values behind it. Without a consensus on values, it is futile to just change the code.

Science fiction author Chen Qiufan says that from a historical and philosophical point of view, humans are now on a path of subject-object dichotomy, longing to overcome it through the eye of AI. This dichotomy not only exists in different genders, races, classes, nations, and cultures of humanity, but also concerns ways of achieving harmonious coexistence between human and other species—a kind of planet-level intelligence that we are counting on to solve all the challenges and ultimate tests we face now and beyond as a civilization.

Daniel Freund, assistant professor of operations management at the

MIT Sloan School of Management, points out that equity issues such as equitable distribution of resources are resolvable. **If technology is used without considering equity, it solves 0% of the problems; conversely, when equity is factored in, it solves 100% of the problems.**

Birat Lekhak, programme and partnerships specialist at UN Women, highlights the **need to build an inclusive, transformative, and responsible innovation ecosystem.** In fact, AI is a product of this ecosystem, and a more inclusive and responsible ecosystem will nurture better AI products. In addition, **we need to empower women in the industry and steer AI toward greater inclusiveness and diversity.**

The panelists agreed that multidisciplinary research between computer scientists and sociologists should be advanced to facilitate algorithmic fairness and address AI-induced biases.

6. Thematic Session #6: Global Digital Governance and Sustainable Enterprise Development

Themed “Global Digital Governance and Sustainable Enterprise

Development,” Thematic Session #6 looks at the risks and challenges faced by enterprises in the context of global digital governance.

Xia Huaxia, vice president and chief scientist at Meituan, notes that data is the “crude oil” of the digital economy and the “blood” of the digital society. **Data governance should not aim to restrict the flow of data, but should guarantee the orderly and free flow of data through governance that ensures compliance in order to promote the development of the digital economy and the rational use of personal information.** As a systematic endeavor, data governance requires diversified management means such as laws, technologies, and markets, as well as the concerted efforts of the international community, countries, businesses, and individuals to ensure the orderly development of the digital market.

Qin Yao, vice president for AI industry development at Huawei, says that the Fourth Industrial Revolution is already here, and the extensive application of AI has accelerated the arrival of an intelligent world. Currently, security and reliability issues have created great challenges for AI application on the ground. Growing

public concern over the risks of AI application has made AI governance a critical issue. **AI should serve to protect the environment, improve human well-being, and address industry issues, and it is therefore important to build open, inclusive, globally governed, and trustworthy AI.** At the same time, it is crucial to develop international multilateral governance mechanisms, improve standards and certification capabilities, build a hierarchical governance framework, and promote inclusive development in building a responsible and trustworthy AI governance system.

Tao Dachen, director of JD Explore Academy and fellow of the Australian Academy of Sciences, points out that the instability of AI systems can easily lead to potential security risks, and the lack of explainability has also limited the otherwise wider AI application and empowerment. How to ensure that user privacy is protected in future AI systems has moved to the center of public attention, demanding immediate action on the development of trustworthy AI. Specifically, **trustworthy AI should be built on four cornerstones: stability, explainability, fairness, and privacy protection.** In the future, trustworthy AI integration will

become an important research discipline, and trustworthiness assessment will be a key step in AI implementation. Overall, continued exploration on trustworthy AI at both the theoretical and practical levels will unleash new waves throughout the AI industry.

According to Song Jiqiang, vice president of Intel Labs and director of Intel Labs China, AI has played a significant role in promoting historical and cultural heritage, protecting wildlife and the natural world, and guarding public health. It is an important technical means to achieve the Sustainable Development Goals. At the same time, the sustainable development of AI is facing challenges ahead: the amount of global data is growing exponentially, and the energy consumption patterns of large-scale AI models are unsustainable.

Laws and regulations, international standards, industry pacts, and corporate self-regulation are the four keys to building an AI governance system for healthy and sustainable development.

Enterprises should be committed to delivering responsibility, inclusiveness, sustainability, and empowerment through technological innovation. They should aim to create a positive impact on society, businesses, and the planet by catalyzing social change, promoting technological inclusiveness, improving health

and security, and contributing to carbon neutrality.

Shen Weixing, dean of the School of Law and dean of the Institute of AI and Law, both at Tsinghua University, also points out that data, as a resource, is the “crude oil” of the digital economy. For data to deliver the most value, he suggests improving relevant laws and regulations and ensuring rational attribution of property rights between users and platforms. An existing model that could be useful is the legal institutionalization of traditional elements (e.g. land). A model of separated ownership and usufruct of data elements could be developed, which means that the user has the ownership of the data, whereas the platform has the usufruct based on its investment of capital, technology, and labor. **At present, the establishment of a reasonable and adequate governance system and framework is required in China’s digital economy to realize a synergy of tripartite empowerment by technology, ethics, and law.** On the one hand, it is necessary to “balance” innovation and regulation. For example, the new *Personal Information Protection Law* has emphasized the need for not only stronger protection of personal information, but also more rational use of it. On the other hand, we need to avoid returning to an

unsustainable path of “development before governance.”

Bill Huang, founder and CEO of CloudMinds Inc., believes that the top three most developed industries in terms of their management and control of information rights are the communications, banking, and publishing industries. Established systems and mechanisms in certain fields may be referenced for insights into data management and control. **It is vital to develop a public computing system that can be used fairly and effectively by all social entities even on a zero-trust basis.** In the future, the “metaverselization” of society as a whole is a challenge we may encounter. China should aim to be a global leader in regulation by establishing end-to-end AI data security regulations, without which challenges will arise.

Xu Jian, chief ecosystem officer and vice president at Horizon Robotics, believes that amid the great opportunities and challenges brought on by digital technologies, a very important area is AI chips in the automotive space. With the rapid advances in intelligent vehicles and autonomous driving, digital AI holds both opportunities and challenges for the automotive industry: the pattern of shared governance has not been fully developed; there

are no established solutions yet to multistakeholder collaboration in the intelligent vehicle industry toward iterative optimization and better data management; and there are no clear definitions of data ownership and of the rules that should be followed in AI training, transmission and security. Data should speed up the flow: edge computing is important for intelligent vehicles—many problems with intelligent driving should be solved on the end side; hardware encryption is another critical topic for safe and reliable autonomous driving.

Meng Qingguo, director of Tsinghua University-China Electronics Corporation Joint Institute for Data Governance Engineering and director of the Center for Internet Governance at Tsinghua University, believes that **a balance between development and security is essential in the era of digital economy**. There exist many risks in data collection, flow, sharing, and application, and the mechanisms and principles behind warrant our consideration. Currently, **there is a comprehensive governance system in which the government, businesses, industry, and civil society all play a role**: the government should lay out industry planning that encourages development while enacting legislation to set standards

on fairness and inclusiveness; enterprises should carry out self-regulation to ensure better compliance, platform openness, fair trade, and security; the industry should set out generally accepted initiatives and launch collective actions on compliance; and civil society should exert an oversight role. On top of the need to protect their own data, multinational enterprises are required to comply with the regulations on cross-border data flows and commercial data protection in both the host and destination countries, and they are perplexed by the decentralized and cliquey regulatory systems. Facing the challenges, **China should be more involved in the establishment of cross-border data governance systems and rules through multilateral international organizations. Greater corporate engagement should be encouraged in the development of national laws and regulations on data, and a more proactive and inclusive strategy should be adopted to address corporate needs regarding cross-border data flows.**

7. Thematic Session #7: Artificial Intelligence and Social Development

Themed “Artificial Intelligence and Social Development,” Thematic Session #7 explores ways to promote responsible and

trustworthy AI for sustainable social development.

Zhu Xufeng, professor and executive associate dean of the School of Public Policy and Management, executive dean of the Institute for Sustainable Development Goals, and deputy director of the Science & Technology Development and Governance Center, all at Tsinghua University, says that the challenges of AI for social development are mainly centered on three areas: **first, technology maturity; second, the two-faced and disruptive nature of its societal impact—a common problem of all technological advances; and third, AI ethics.** All these aspects need further categorization and refinement. Given international political factors, we need to consider, in any technology transfer efforts, not only the willingness of technologically advanced countries to transfer their technologies to less competitive powers, but also the willingness of the latter to take them in. For example, distrust may lead the latter to turn down AI technologies that involve big data or financial services, out of the fear of potentially handing over their own financial systems.

Christoph Lütge, director of the TUM Institute for Ethics in Artificial Intelligence and academic member of I-AIIG, points out

that nuclear technology is generally considered to be highly risky in previous philosophies of technology: a nuclear war could wipe out the entire human race. Therefore, there is a consensus among countries on the perils of the technology. Unfortunately, there is no such clear understanding of the impact of AI. **It is like a black box; we are concerned, yet we don't know at what point its risks will materialize.** Technology transfer is supposed to be a process of mutual learning and enrichment. And such interaction between technical and non-technical professionals as well as between different cultures and civilizations is essential for the concerns and resistance of countries to be heard regarding the use of AI. Therefore, there is a need for more debate on ethics and global governance.

Zeng Yi, co-director of the China-UK Research Centre for AI Ethics and Governance at the Institute of Automation, Chinese Academy of Sciences, and chief scientist at I-AIIG, believes that building trustworthy AI requires not only a technological approach, but also a societal one. As we debate on public trust of AI, it is necessary to explore the trust between different regions and cultures. Whether divergences in cultural values will result in the

distrust of one country in the AI products developed in another country is an issue that merits our consideration. Mutual cultural trust and respect are essential in technology transfer, where a product or platform needs to adapt to a new cultural context and understand the needs and concerns of the destination country so that it can win its trust and be put into use. In addition, infrastructure sharing is somehow more important than product sharing, as it allows the recipient to create products and services that best serve their national conditions, human resources, and needs.

Tokuchi Tatsuhito, executive director at the Center for Industrial Development and Environmental Governance at Tsinghua University and former managing director at CITIC Securities Co., Ltd, says that if emerging technologies such as AI are not well controlled in some developing countries, they can easily spread, resulting in conflicts with neighboring powers. He believes that **well-managed export of emerging technologies to developing countries is good for sustainable development**, and international organizations need to set up relevant management and control systems, which could otherwise lead to tragic consequences for the

region or even humankind as a whole. Long-term efforts and consensus among countries are needed in addressing the disparities between technologically advanced and less competitive powers. More importantly, **it is crucial to invest more in education so that children and young people can receive equal education on AI.**

Jia Xiaofeng, director of the Data Management Department at Beijing Big Data Center, touches on the nature of AI, its operational mechanisms, and the changes it brings to industry chains. AI has played a significant role in promoting sustainable development by enabling greater information symmetry and more accurate resource matching, and by serving as a new generation productive force for industry, society, research, and military. **AI is not about doing things better than before, but about fundamentally transforming the current governance system, and thus the way we live and produce, and even future livelihoods.** Promoting responsible and trustworthy AI requires consensus on three things. First, **the international community should develop mechanisms that allow for dynamic consensus on basic AI guidelines and laws,** including institutional systems, policies, energy consumption, technical routes, etc. Second, the

development process of AI from basic theory to engineering and then to industrialization is not a single-threaded one. Therefore, **it is necessary to promote interdisciplinary crossover and to facilitate external diffusion in addition to internal incubation.** Third, **it is important to guard against future systemic risks.** It is not only necessary to solve the problems we face today, but also to make further digital incremental layouts for the coming five to ten years and beyond.

Alexander Kriebitz, research associate at the TUM Institute for Ethics in Artificial Intelligence, believes that AI can significantly improve regional productivity. For example, parts of Africa that were not industrialized have taken the lead in digitization and made rapid progress. Therefore, **we need to find a mission and vision for AI and identify the Sustainable Development Goals that are expected to be achieved with the help of AI.** AI should be used with greater caution in government, as it opens up legitimacy issues. It is important to ensure that government actions are based on popular consent and are constitutional. At the same time, **AI explainability needs to be improved, otherwise it will be difficult to ensure public trust in the technology.**

The panelists have always believed that while AI is empowering businesses from every industry, it has also triggered security, privacy, and fairness concerns and challenges. To address these problems, it is necessary to strengthen global collaboration, introduce relevant policies and regulations, and develop a balanced and inclusive AI governance system, so as to steer the sustainable and responsible development of AI for the betterment of human society.

IV. Next Step

The International AI Cooperation and Governance Forum 2022 will be held in the late of the year, please stay tuned.

Forum Website: <https://www.tsinghuaaiforum.org/>

I-AIIG Website: <https://aiig.tsinghua.edu.cn>

Contact: i-aiig@tsinghua.edu.cn