Issue 14 · March 24, 2024

China Development Forum 2024

The Continuous Development of China

Symposium on AI Development and Governance (Panel Discussion II)

Hosted by the Development Research Centre of the State Council (DRC) and organized by the China Development Research Foundation (CDRF), the China Development Forum (CDF) 2024 was held at Diaoyutai State Guesthouse in Beijing from March 24th to March 25th. On the afternoon of March 24, the "Symposium on AI Development and Governance" took place. "Panel Discussion II" featured five speakers: Xue Lan, Dean, Schwarzman College, Tsinghua University; Andrew Forrest, Executive Chairman, Fortescue; Zhang Yaqin, Dean, Institute for AI Industry Research (AIR), Tsinghua University; Academician, Chinese Academy of Engineering (CAE); Robert Moritz, Global Chairman, PwC; and Li Cheng, Professor, The University of Hong Kong; Director, Centre on Contemporary China and the World; and Scott Kennedy, Senior Adviser and Trustee Chair in Chinese Business and Economics, Center for Strategic and International Studies (CSIS). The symposium was chaired by Zhang Shunxi, vice president of the DRC.

In addressing AI risk prevention and governance, Xue Lan

identified five challenges. The first challenge is the issue of inconsistency. AI technology evolves swiftly, yet the adaptation of governance systems lags due to the extensive deliberation and scrutiny required for each law and regulation. The second challenge is the issue of information asymmetry. Investigations reveal blind spots in both the government, as the governing body, and enterprises, as the regulated entities. Enterprises are often unaware of the government's primary concerns in governance, while the government may not fully grasp the risks technology can pose. This information asymmetry between the two can sometimes result in mutual unawareness. The third challenge is the cost asymmetry in risk regulation. The costs of misusing or abusing AI are relatively low, while the costs to safeguard against technological risks far exceed the potential damage, leading to high governance expenses. The fourth challenge pertains to the global governance system. Global governance typically involves a complex array of mechanisms, and no single country, international organization, or enterprise can tackle these issues alone, as exemplified by AI. Certain fields may attract professional bodies, international organizations, enterprises, and other entities with vested interests, conditions, and capabilities, all eager to engage in governance. These organizations might lack affiliation, experience overlap, or even encounter contradictions, complicating the establishment of a global governance system. The fifth challenge involves geopolitical issues. The fierce scientific and technological rivalry between China and the United States raises issues of technological suppression due to geopolitics. Thus, China and the

United States must collaborate on AI governance.

To address the issue, Xue Lan suggested that firstly, bolstering security-focused technology research and development, particularly through international cooperation, is essential; secondly, agile governance should be promoted; thirdly, enterprises should be encouraged to self-regulate internally; fourthly, global governance needs to be reinforced; and fifthly, dismantling "small yard, high fence" and enhancing China-US cooperation in AI is crucial.

Andrew Forrest asserted that AI was intimately linked to human survival; it's one of humanity's greatest inventions, yet also a perilous innovation. Whether considering the standpoint of China, the United States, or the global community, AI must be regulated and humans must retain the capacity to control it. AI essentially stems from concepts shared by everyone online, and it evolves at an extraordinary pace. If left unregulated, human intelligence could be surpassed by superintelligent species that are rapidly improving and iterating, presenting a grave danger to the entire world. Should an error occur in AI application, a consensus-driven approach is needed to allow human intervention without catastrophic consequences.

Zhang Yaqin concisely outlined the evolutionary trends of large-scale AI models. Firstly, new intelligence will emerge that is multi-modal, multi-scale, and cross-modal; secondly, within the next five years, significant breakthroughs in overall architecture are anticipated, adhering to the Scaling Law, though not necessarily via the transformer model structure; thirdly, intelligence is

progressively shifting towards the edge with devices like "AI smartphones" and "AI PCs"; fourthly, the capability for autonomous intelligence in defining tasks, planning routes, self-improvement, and self-coding will be realized; fifthly, intelligence based on information is gradually transitioning into the physical realm; and sixthly, there will be a connection between large models and biological intelligence.

Zhang Yaqin contends that in the coming five years, the expansive deployment of AI technology in key domains will introduce three risks: the risk of misinformation in the digital sphere, including the spread of errors and falsehoods; the issue of large model illusions, which will amplify as the use of information intelligence broadens to encompass physical and biological intelligence; and the hazards associated with large AI models integrating with economic, financial, military, and power grid systems.

Zhang Yaqin offered five long-term recommendations for mitigating these risks. The first is to label AI-generated digital humans and other intelligent entities similarly to advertisements. The second is to create a mapping and registration system. Clarify that the robot, as a subordinate entity, must be linked to a responsible party, which can be a legal entity like an individual or a company, allowing for accountability if issues arise. The third is to set up a tiered oversight system and mechanism. Implement regulation for large models across various domains, including physical and biological systems, at differentiated levels. Fourth, boost investment in large model risk research and advocate for

collaborative engagement from governments, scientists, tech practitioners, and entrepreneurs to achieve development alongside governance. Fifth, establish a red line for AI development. Currently, the Beijing International Dialogue on AI Security has culminated in the Beijing International Consensus on AI Security, proposing a red line for AI development. All nations should collaborate in the realm of global intelligence to collectively address AI redlines as societal risks.

Robert Moritz primarily discussed the governance of AI. The first point is risk, encompassing macro, micro, and existential dangers. Experimentation and learning are essential; it's important to keep abreast of developments, capitalize on AI opportunities, expand advantages, avert risks, and drive transformative development.

Robert Moritz posited that alongside AI regulation at governmental and societal levels, corporate management practices must also evolve, including shifts in values, incentive structures, behaviors, and culture, to more comprehensively consider the impacts of AI. Hence, it's essential to reform enterprise management systems, expedite decision-making, and address shortcomings. To maximize strengths and minimize weaknesses, change should not be confined to a single organization but rather orchestrated across the ecosystem, uniting multiple businesses to ensure smooth plan execution. A stronger connection will lead to more favorable development.

Li Cheng remarked that AI has profoundly changed everything except the way people think. AI is affecting society with

unprecedented speed, scale, scope, and intensity, but there is a lack of readiness for it, including in areas of governance and employment. It must be understood that AI is merely a tool, and its usage and effects are not yet fully certain. However, it's crucial to stay abreast of AI developments, reflect deeply, and devise strategies to mitigate any negative impacts. As leading powers, China and the United States should urgently consider the future applications of AI, including data access and usage scope, and progress collaboratively.

Scott Kennedy outlined three current challenges confronting AI. The first challenge is the security of AI applications, including concerns over autonomous driving and large language models. The second challenge involves AI technology and its applications, where international circulation is still restricted. In the upcoming period, the question is how to gain the initiative in the three core elements of AI—data, algorithms, and computing power—to seize developmental opportunities. The third challenge is AI governance, encompassing three dimensions: multilateral, bilateral, and unilateral. At the multilateral level, there's a need to bolster multilateral governance for safe, trustworthy AI and sustainable growth. At the bilateral level, achieving consensus and cooperation to develop framework agreements for AI governance is crucial. On the unilateral front, numerous actions are possible, such as China and the United States crafting comprehensive laws and frameworks to regulate data privacy and security, dismantling "firewalls" to boost AI development and improve people's lives; China should also increase its participation in setting international

AI standards, thereby expanding its influence and voice, presenting itself more effectively to the world, addressing issues, and fostering shared progress.

(China Development Press Authors: Liu Changjie, Luo Rensheng, Xu Jing; Reviewer: Yang Liangmin)

--Background Information--

Under the mandate 'Engaging with the world for common prosperity', China Development Forum (CDF) serves as an important platform for Chinese government to carry out candid exchanges and discussions with leaders of global businesses and international organizations as well as foreign and Chinese scholars. Initiated in 2000, CDF has made remarkable contributions for the policy exchange and international collaborations between China and the world.

--Media Contact—

CDRF

Guo Silu 18666028168/64255855-8014

Shi Yafan 13810361966/64255855-8223

Xia Tian 18801375838/64255855-8086

Shi Wanjing 18801090391/64255855-8090

Qiu Kaixian 18301078627/64255855-8103