

China Development Forum 2024 The Continuous Development of China

Symposium on AI Development and Governance (Keynote Speech 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. The "Symposium on AI Development and Governance" took place on the afternoon of March 24, featuring Zhang Hongjiang, Founder and Founding Chairman, Beijing Academy of Artificial Intelligence, as the speaker for "Keynote Speech II". The symposium was chaired by Zhang Shunxi, vice president of the DRC.

Zhang Hongjiang suggested that the future trajectory of today's popular multi-modal large models should extend beyond just generating videos, editing footage, making movies, or producing TV series. From a technical perspective, these models could act as machine cognition to discern the surrounding world and equip autonomous driving systems, effectively evolving current information and model systems into the dynamic systems of tomorrow.

Zhang Hongjiang found the prospect of multi-modal large models, especially within motion systems, equipping robots with a brain to be the most exhilarating aspect. For instance, if a robot were instructed to pick an extinct animal from a pile of toys on a table, it could successfully identify and retrieve an extinct dinosaur among several known animals like tigers, lions, and birds through reasoning and recognition. Such a task was beyond the capabilities of robots in the past. Previously, you could instruct a robot on what to catch, but it couldn't handle an abstract concept. Likewise, tell a contemporary robot that it is thirsty, and the robot will pick up a bottle of water from a collection of items. These demonstrations illustrate that when equipped with large multimodal models, robots are not just following prior instructions but can interpret the meaning behind them. This is the prototype of the autonomous robots envisioned for the future. Multimodal large models are already capable of astonishing us today.

This is actually the culmination of AI's continuous evolution over the past 70 years. AI has experienced three waves of development. The third wave has been marked by the surge of deep learning over the past decade. The advent of large models coincided with the trough of the third wave in 2020, with ChatGPT-3 marking a turning point that led to Sora and a series of large-scale models in China. Over the past ten months, there's been swift progression from language and multimodal models to visual and, eventually, larger-scale future models. What drives this evolution? A crucial factor is that AI today has transcended being merely an algorithm to becoming a system. Modern models are not just large

in scale, but also versatile. The impetus for this is the the Scaling Law. This scaling effect enables the resolution of numerous problems that were previously unsolvable. Using the language model as an example, with only billions of data points, it could only tackle certain issues in traditional natural language processing. Yet, once it surpasses 500 billion data points, virtually all natural language problems can be effortlessly resolved, illustrating what is known as scaling capability.

What is the core behind large models? Zhang Hongjiang considers it a new operating system. In the traditional PC era, output was generated through CPU computations. Today, the heart of large model computing has shifted from the CPU to the GPU, signifying a new operating system. Currently, all internet platform companies are racing to develop large models, driven by the underlying notion that without a large model, they won't remain platform companies in the future. Observing the progress over recent years, particularly the last 18 months, it's evident that the new Moore's Law posits that with each one to two generational leaps in model capability, training costs drop to a quarter of their previous amount every 18 months, while the inference cost of a model decreases to one-tenth every four months. This new iteration of Moore's Law will drive the swift proliferation, advancement, and deployment of large models. Another catalyst is Nvidia's stock performance; over the past 12 months, it has skyrocketed, positioning it as one of the world's top three most valuable companies. The entire large model industry chain is expanding and evolving at a rapid pace. Large models will invigorate our software

tools, enhance our daily lives, and enrich our work.

Of course, there are some risks associated with the future development of AGI that could even lead to human extinction. Geoffrey Hinton took a keen retrospective on AI and neural network developments since last March, concluding that digital computing has surpassed biological computing in efficiency, as digital systems can learn from each other and transfer their accumulated knowledge to subsequent systems. This implies that AGI is on the cusp of surpassing human intelligence in all aspects.

Historically, human learning capacity has followed a relatively linear trajectory, whereas machine learning ability charts an exponential curve. Over time, as these two lines intersect, machine learning capability will surpass that of humans. Thus, the jobs that machines displace will be replaced by new machines rather than by humans. So, will human jobs vanish entirely? Will the majority face unemployment? In reality, unemployment isn't the worst outcome. The most alarming prospect is that if the system becomes capable of self-improvement, it could pose profound risks to humanity. To avert the potential global catastrophic consequences of AI, we need to take decisive action now and bolster international collaboration across technical, legislative, and governance domains.

Zhang Hongjiang observed that AI has entered a new developmental phase with large models, symbolizing the fourth technological revolution (following the agricultural, industrial, and information revolutions). This will lead to significant efficiency gains, offer considerable convenience to our lives, and generate immense

value and numerous new industries. However, we must also recognize the global and potentially catastrophic consequences of AI, which could pose an existential threat to humanity. To prevent such risks, we must establish clear boundaries and enhance governance mechanisms; concurrently, we need to advance security technologies to prevent AI from overstepping these red lines. For this to occur, it's crucial to sustain and intensify security collaboration among the international scientific and policy communities to avoid such a catastrophe.

(China Development Press Authors: Liu Changjie; 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



中国发展研究基金会
China Development Research
Foundation

中国北京东城区安定门外大街138号
皇城国际中心A座15层

邮编: 100011
电话: 86-10-64255855
传真: 86-10-64255855-8100
网址: www.cdrf.org.cn
电邮: cdrf@cdrf.org.cn

Floor15, Tower a, Imperial International Center, No.138
Andingmen Wai Avenue, Dongcheng District, Beijing, 100011, China

Tel: 86-10-64255855
Fax: 86-10-64255855-8100
Website: www.cdrf.org.cn
E-mail: cdrf@cdrf.org.cn