EDITORIAL

Artificial intelligence for science—bridging data to wisdom

Yongjun Xu,[1,*] Fei Wang,[1] Zhulin An,[1] Qi Wang,[1] and Zhao Zhang[1,*]

[1]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

*Correspondence: xyj@ict.ac.cn (Y.X.); zhangzhao2021@ict.ac.cn (Z.Z.)

Received: October 10, 2023; Accepted: October 17, 2023; Published Online: October 18, 2023; https://doi.org/10.1016/j.xinn.2023.100525

Citation: Xu Y., Wang F., An Z., et al., (2023). Artificial intelligence for science—bridging data to wisdom. The Innovation **4(6)**, 100525.

INTRODUCTION

Throughout the scientific discoveries of human history, research paradigms have undergone profound changes.[1] As shown in Figure 1, from the empirical paradigm, which relies on experimental observation, to the theoretical paradigm, which is based on theoretical deduction, and further to the computational paradigm, which is associated with simulation and emulation, and finally to the data-driven paradigm, which is grounded in human-machine-object integration, every transformation of scientific paradigm has brought about a major leap in science and technology.[2] In recent years, with the increasing volume of scientific data and the continuous progress in information processing and knowledge discovery technology, artificial intelligence (AI) technology has been gradually applied to various scientific fields. In 2021, Xu et al.[3] proposed the AI paradigm (AI4Science) in *The Innovation*, according to which AI can make breakthroughs in the application of multiple scientific fields and solve problems the traditional paradigms could not solve.

From an article published in *Science* in 2015 proposing that AI passed the Turing test,[4] to AlphaGo's success in defeating humans in the game of Go in 2016, to DeepMind's AI-based nuclear fusion control method, to AlphaCode's achievement of competitive programming skill, and then to Huawei's successful resolution of key challenges in global weather forecasting with its PanGu AI model, all of these demonstrate the significant role that AI is playing in various scientific fields and the gradual achievement of landmark results, bringing about revolutionary progress. In August 2023, a review paper[5] was published in *Nature* discussing the application value of AI in scientific discovery across multiple fields. To sum up, it can be seen that AI technology is gradually changing the way of scientific research, providing important technical support and new development prospects

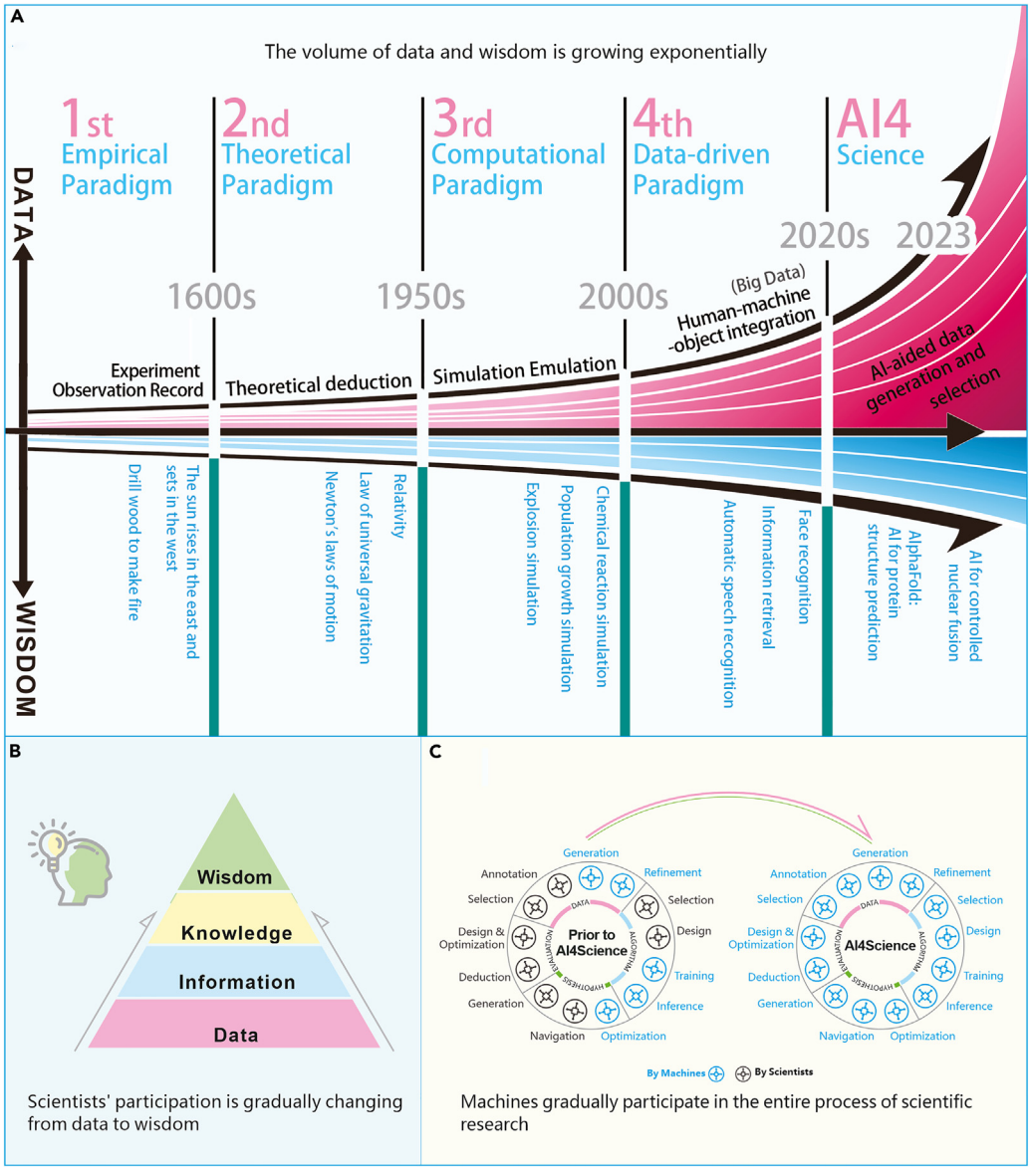

Figure 1. AI4Science: bridging data to wisdom

for further development of various fields. A new scientific revolution seems imminent. So what are the main features of AI4Science? And what drives science to enter the AI paradigm?

CURRENT STATUS OF AI4Science

According to our observation, in the ever changing landscape of the human research paradigms, the generation of scientific data has undergone rapid evolution. In the past, researchers obtained data primarily through real-life experiences or scientific experiments. Data are very limited and subject to constraints imposed by time and resources. With the advancement of technology, computer simulations have become an important method of data generation. Simulation allows researchers to create virtual environments for data generation, promoting the rapid development of science, but it has limitations in terms of data quality and diversity. In recent years, the introduction of high-performance computing, big data, and AI technologies has provided efficient and autonomous methods of data generation. For example, reinforcement learning generates data by enabling intelligent agents to interact with their environment, which has promoted intelligent research in various fields. In the aforementioned processes, both the scale of data and the volume of knowledge have continuously increased millions of times. It helps researchers gain a more comprehensive understanding of the behaviors and developmental patterns of complex systems. For instance, the Five-hundred-meter Aperture Spherical radio Telescope (FAST) generates a staggering 500 TB of raw data every day. The continuous development of intelligence computing empowers researchers to tackle increasingly complex scenarios and issues, with data volumes reaching the petabyte, exabyte, or even zettabyte level. In a certain sense, the volume of scientific data and the speed of scientific discovery directly determine the scientific paradigms of different eras.

Despite the increases in data scale, knowledge volume, and computational capabilities, which have provided strong support for scientific research, they have also brought new challenges. Specialization within disciplines has led to increasingly specialized research, while the demand for interdisciplinary collaboration has been steadily rising. The transition from data to wisdom has become progressively more challenging. In this context, the innovation of scientific paradigm has become the key to problem solving. AI technology contributes to achieving reasoning in more complex systems and leads to the emergence of more interconnected and cross-disciplinary research directions. Different from previous scientific paradigms, machine conjecture, knowledge creation, and intelligence emergence are key characteristics of AI4Science. For instance, AlphaFold has achieved promising results in the problem of protein structure prediction. NASA has released the first geospatial AI foundational model to support the resolution of complex geoscience issues. Google has used large language models to create AI doctors, which are now on par with human doctors in answering medical questions. DeepMind has used AI to guide human intuition in discovering mathematical theorems. In conclusion, paradigm innovation based on AI technology is the key to addressing current challenges. It has the potential to revolutionize research methodologies, enabling a more effective transition from "knowledge under the data" to "wisdom from the data" and thus enabling researchers to better address increasingly complex scientific problems.

FUTURE OUTLOOK

With the gradual increase in the volume of data, as well as the continuous progress of information use and knowledge discovery technology, the use of AI tools to extract crucial knowledge from extensive datasets will progressively emerge as a vital avenue to foster the continued advancement of scientific research and will have a far-reaching impact on various classic and important scientific research. In the future, scientists will increasingly depend on data collection, big data storage capabilities, and intelligent processing capabilities. And the core technology will also develop in the direction of automation as well as intelligence, changing the scientific data processing from human in the bypass to the scientific discovery of human in the loop. Overall, the future paradigm of scientific research will be changed, and scientific discovery based on AI will gradually become an important route of scientific research.

In this case, more and more professional intelligent tools need to be created in all aspects of scientific research. These tools will be used in a wide range of disciplines, generating a series of intelligent research platforms with different characteristics. They will be able to analyze scientific problems from a more professional perspective, discover more core knowledge systems, create more perfect research routes, and further solve critical and difficult problems in various scientific fields. By using intelligent tools, scientists will be liberated from the burden of generating massive amounts of experiments, enabling them to expedite the development of scientific ideas, the design and progress of innovative technologies, and the discovery and emergence of specialized knowledge. This will propel scientific development into a new technological stage, creating a new paradigm for scientific research that leverages data to explore information, generate knowledge, cultivate wisdom, and seamlessly integrates with scientists' innovative ideas through AI technology. In conclusion, the essence of AI4Science is to help scientists bridge scientific data to scientific discovery.

REFERENCES

1. Kuhn, T.S. (1996). The Structure of Scientific Revolutions (University of Chicago Press).
2. Hey, T., Tansley, S., and Tolle, K. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery (Microsoft Research).
3. Xu, Y., Liu, X., Cao, X., et al. (2021). Artificial Intelligence: A powerful paradigm for scientific research. Innovation ***2***, 100179.
4. Lake, B.M., Salakhutdinov, R., and Tenenbaum, J.B. (2015). Human-level concept learning through probabilistic program induction. Science ***350***, 1332–1338.
5. Wang, H., Fu, T., Du, Y., et al. (2023). Scientific discovery in the age of artificial intelligence. Nature ***620***, 47–60.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under grants 62372430 and 62206266 and the Youth Innovation Promotion Association CAS under grant 2023112.

DECLARATION OF INTERESTS

The authors declare no competing interests.