GUEST EDITORIAL

National Science Open
3: 20240008, 2024
https://doi.org/10.1360/nso/20240008

Chemistry

Special Topic: AI for Chemistry

Artificial intelligence is empowering chemistry research

Shanshan Wang[1,2] & Jin Zhang[1,3*]

[1]*School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China;*
[2]*College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China;*
[3]*School of Material Science and Engineering, Peking University, Beijing 100871, China*

*Corresponding author (email: jinzhang@pku.edu.cn)

Received 13 March 2024; Revised 13 March 2024; Accepted 13 March 2024; Published online 20 March 2024

In recent years, the rapid development of big data and artificial intelligence (AI) is increasingly changing the research paradigms of science and technology. AlphaFold 2 has predicted the three-dimensional structures of 214 million proteins in more than one million species, covering almost all known proteins on Earth. Deep reinforcement learning has realized the magnetic control of the tokamak plasma, which requires high-dimensional, high-frequency, closed-loop controls together with a bunch of physical and operational constraints. AI also automatedly designed an industrial-scale central processing unit (CPU) within 5 h by exploring a search space of the unprecedented size of $10^{10^{540}}$. All these examples show the enormous potential of AI to revolutionize the way research is done in the scientific field.

Chemistry, as a basic science that studies the composition, structure, properties of substances, and the transformations they undergo, has long suffered with huge parameter space, large span of dimensions, and low experimental efficiency due to traditional trial and error methods. Luckily, in recent years, with the deep intersection of AI and chemistry, these problems are gradually being solved. High-precision, high-efficiency, multi-scale chemical prediction models have been developed. Large model-driven, human-machine collaborative intelligent chemical laboratories have been constructed. Empowered by AI, researchers have realized more efficient optimization of process parameters and a deeper understanding of the structure-activity relationship, thus accelerating the discovery of breakthrough theoretical laws and the synthesis of innovative substances.

Here we organize a special topic on "AI for Chemistry", which includes seven high-quality papers covering the latest research and reviews of AI-enabled structure-property prediction models and automated robotic experimental systems in the fields of organic chemistry, energy chemistry, biochemistry, and nanomaterial-related systems.

Machine learning algorithms excel at constructing theoretical models that balance efficiency and accuracy and solve multi-scale simulation problems in complex chemical systems. Gong *et al*. [1] reported machine learning potential, which combined the basin-hopping Monte Carlo algorithm, deep learning methods, molecular dynamics simulations, and density functional theory calculations, to simulate nanoparticles of

varying sizes across a wide temperature range, particularly in the nanometer scale. Yu *et al*. [2] developed a machine learning model based on the artificial neural networks embedded graph attention network, which used the micro-scale and the macro-scale simulations as the training datasets, to realize the multi-scale simulation of the electrical resistance-strain response of the carbon nanotubes/polydimethylsiloxane composites. Li *et al*. [3] proposed a novel strategy, which combines molecular docking with machine learning methods, to predict the protein-ligand complex structures. This strategy inherits the powerful predictive ability from the machine learning methods and the physical constraints advantages from the traditional molecular docking approaches so that both the success rate and the accuracy of the prediction results have been significantly improved.

Energy chemistry related to the development of catalysts and batteries is one of the most widely applied areas of AI today. Zhang *et al*. [4] developed a multi-source semantic knowledge graph framework, which can capture the latent knowledge of reactant-catalyst-product relationships and provide accurate recommendations on potential catalysts for targeted reactions. Wang *et al*. [5] applied first principles and machine learning strategies to study the catalytic activities of 4d/5d period transition metal single-atom catalysts (SACs) embedded on MBene substrates for oxygen reduction and evolution reactions (ORR/OER). The machine learning model involved 10 element features of SACs and displays fast and accurate identification of the excellent ORR and OER electrocatalysts. Chen *et al*. [6] systematically reviewed the AI application in lithium battery chemistry from the aspects of the electrolyte design, electrode interfacial simulations, lithium dendrite growth, and battery life predictions. The comprehensive review shows the great potential of machine learning methods in the data-driven experimental optimization, mechanism revealing the ability of the lithium dendrite growth and the prediction power of the battery lifespan, which facilitates the development of more efficient, safer, and longer-life batteries.

AI can also alleviate chemists' manual and labor-intensive tasks and allow them to concentrate their efforts on scientifically impactful pursuits. Details include the construction of high-throughput automation platforms to realize efficient and consistent experiments, as well as the development of machine learning algorithms to automatically analyze and mine the large-volume data generated by automated experiments, thereby revealing the deep-seated laws. Liu *et al*. [7] reviewed the recent progress in the transformation of organic chemistry research paradigms inspired by AI. Machine learning models together with autonomous robotic systems dramatically increase the planning ability for complex molecular synthesis and accelerate the exploration pace due to the unprecedented speed and precision of automated experiments.

As the space of this special topic is limited, we cannot list all the recent progress made in the field of AI for Chemistry. However, we believe that this topic will inspire researchers to integrate AI with chemistry, drive a paradigm shift in chemistry research, and promote discoveries and team building in cross-disciplinary areas. We would like to thank all the authors who have contributed high-quality peer-reviewed articles to this special topic. We are also grateful to the deputy editors in chemistry who invited these papers, as well as the editorial and production staff of the *National Science Open* for their high-quality assistance.

References

1 Gong FQ, Xiong K, Cheng J. Constructing machine learning potential for metal nanoparticles of varying sizes via basin-hoping Monte Carlo and active learning. *Natl Sci Open* 2024; **3**: 20230088.

2 Yu L, Zhi C, Sun Z, *et al.* A machine-learning-enabled approach for bridging multiscale simulations of CNTs/PDMS composites. *Natl Sci Open* 2024; **3**: 20230055.

3 Li Y, Lin H, Yang H, *et al.* Synergistic application of docking and machine learning for improved binding pose. *Natl Sci Open* 2024; **3**: 20230058.

4 Zhang Z, Ma S, Zheng S, *et al.* Semantic knowledge graph as a companion for catalyst recommendation. *Natl Sci Open* 2024; **3**: 20230040.

5 Wang E, Wang G, Zhou J, *et al.* MBenes-supported single-atom catalysts for oxygen reduction and oxygen evolution reactions by first-principles study and machine learning. *Natl Sci Open* 2024; **3**: 20230043.

6 Chen Y, Liu Y, He Z, *et al.* Artificial intelligence for the understanding of electrolyte chemistry and electrode interface in lithium battery. *Natl Sci Open* 2023; **2**: 20230039.

7 Liu C, Chen Y, Mo F. Transforming organic chemistry research paradigms: Moving from manual efforts to the intersection of automation and artificial intelligence. *Natl Sci Open* 2023; **2**: 20230037.