



Contents lists available at ScienceDirect

Journal of Materiomics

journal homepage: www.journals.elsevier.com/journal-of-materiomics/

Generative artificial intelligence and its applications in materials science: Current situation and future perspectives



Yue Liu ^{a,d}, Zhengwei Yang ^a, Zhenyao Yu ^a, Zitu Liu ^a, Dahui Liu ^a, Hailong Lin ^b,
Mingqing Li ^b, Shuchang Ma ^a, Maxim Avdeev ^{e,f}, Siqi Shi ^{b,c,*}

^a School of Computer Engineering and Science, Shanghai University, Shanghai, 200444, China

^b State Key Laboratory of Advanced Special Steel, School of Materials Science and Engineering, Shanghai University, Shanghai, 200444, China

^c Materials Genome Institute, Shanghai University, Shanghai, 200444, China

^d Shanghai Engineering Research Center of Intelligent Computing System, Shanghai, 200444, China

^e Australian Nuclear Science and Technology Organisation, Sydney, 2232, Australia

^f School of Chemistry, The University of Sydney, Sydney, 2006, Australia

ARTICLE INFO

Article history:

Received 2 May 2023

Received in revised form

10 May 2023

Accepted 12 May 2023

Available online 25 May 2023

Keywords:

Machine learning

Artificial intelligence

Generative artificial intelligence

Materials science

Novel materials discovery

Deep learning

ABSTRACT

Generative Artificial Intelligence (GAI) is attracting the increasing attention of materials community for its excellent capability of generating required contents. With the introduction of Prompt paradigm and reinforcement learning from human feedback (RLHF), GAI shifts from the task-specific to general pattern gradually, enabling to tackle multiple complicated tasks involved in resolving the structure-activity relationships. Here, we review the development status of GAI comprehensively and analyze pros and cons of various generative models in the view of methodology. The applications of task-specific generative models involving materials inverse design and data augmentation are also dissected. Taking ChatGPT as an example, we explore the potential applications of general GAI in generating multiple materials content, solving differential equation as well as querying materials FAQs. Furthermore, we summarize six challenges encountered for the use of GAI in materials science and provide the corresponding solutions. This work paves the way for providing effective and explainable materials data generation and analysis approaches to accelerate the materials research and development.

© 2023 The Authors. Published by Elsevier B.V. on behalf of The Chinese Ceramic Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

1. Introduction	799
2. Various GAI methods and their applications to materials science	799
2.1. Generative Adversarial Network	799
2.2. Variational autoencoder	803
2.3. Diffusion model	803
2.4. Flow model	804
2.5. General GAI	805
3. Future applications of general GAI in materials science: ChatGPT example	808
3.1. Generating materials data	808
3.2. Solving differential equation	808
3.3. Querying materials FAQs	810
4. Analysis of problems of GAI in materials science	812
4.1. High-quality data and domain knowledge	812
4.2. Generalization ability	812

* Corresponding author. State Key Laboratory of Advanced Special Steel, School of Materials Science and Engineering, Shanghai University, Shanghai 200444, China.

E-mail address: sqshi@shu.edu.cn (S. Shi).

Peer review under responsibility of The Chinese Ceramic Society.

4.3. Interpretability and credibility	813
4.4. Usability	813
4.5. Resource cost	813
4.6. Security	813
5. Conclusion	814
Declaration of competing interest	814
Acknowledgment	814
References	814

1. Introduction

Data-driven machine learning (ML) brought the possibility of the “4th paradigm” of materials research and development due to the effective ability of ML to uncover the structure-activity relationships in materials data [1]. However, the paradigm shift is hindered because of slow progress in data quality governance and the lack of the guidance for using domain knowledge in combination with data-driven analysis, which can be summarized as three key issues, *i.e.*, high dimensionality of feature space *vs.* small sample, model accuracy *vs.* usability, and ML results *vs.* domain knowledge [2]. Embedding domain knowledge into models with generative ability would resolve the above-mentioned issues and enable such models to mine structure-activity relationships more accurately.

Generative Artificial Intelligence (GAI), an unsupervised or partially supervised ML framework [3], enables the input samples to be semantically edited or disassembled through disturber of the input distribution or generation of the distribution for fitting the posterior distribution of the target samples. Such operations can assist in analyzing the inherent relationship between features while generating target samples accurately, which has been widely employed in image and video generation, sequence modeling, speech enhancement, *etc.* [4,5]. For materials science, the target materials data can be edited through embedding physical rules and operations (*e.g.*, translation and rotation) into the generation process of GAI, which has been applied to material performance prediction and new material discovery [6–8]. For example, Hoogeboom *et al.* [9] introduced different Euclidian groups into the generation process of GAI, to improve quality of generated samples. Zhao *et al.* [10] modified the generation process by the description of crystal space information, to efficiently generate novel materials with high crystal structural diversity. These studies illustrate that GAI is conducive to the deep coupling of AI for Science (AI4Science) [11] with materials science, thus accelerating the progress.

Nowadays, researchers tend to train complicated generative models on large datasets, leveraging the Prompt paradigm [12] to accomplish the few-shot learning [13] even zero-shot learning [14]. Then, through the combination of reinforcement learning from human feedback (RLHF) and such models (*e.g.*, ChatGPT) [15], domain knowledge can participate in the model training process in a highly coupled way, which drives the transfer of the paradigm from “Fitting-Generation” to “Pretraining-Prompting-Generation”. There have been studies that demonstrated high accuracy of the content generation in multimodal data generation and analysis, multi-context reasoning, and code design [16,17], which is promising for achieving realization of the general GAI. These models can be employed to modify and generate materials data of interest through the multimodal data generation and domain-knowledge-embedded generation patterns, thereby breaking the barriers of different scales, systems, and types of materials. However, the current development of general GAI still suffers from some key

issues, *e.g.*, high training and maintenance costs, high-quality data sparseness, poor domain knowledge integration, model interpretability and credibility, resource sharing, security, *etc.*

In this paper, we review the status of GAI development systematically and summarize the cons and pros of different GAI models and scope of their applications. Then, the studies of using GAI for materials science are reviewed and the future applications of GAI in materials science are discussed, aiming to aid researchers to rapidly select the proper generative models. Following this, we perform trials, taking ChatGPT as an example, to explore how general GAI can solve the high-profile materials issues, *e.g.*, novel material data generation, the solution of differential equations and query of materials FAQs. Finally, the challenges of the development of GAI for materials science are analyzed and discussed in detail, such as the requirements of high-quality data and domain knowledge, the improvement of model generalization, interpretability and credibility, usability, security, and resource cost.

The remainder of this paper is organized as follows: Section 2 summarizes the development process of various GAI models from the perspective of specific tasks and general types and introduces their applications in materials science. Sections 3 explores possible applications of general GAI for dealing with the domain key issues. The challenges of development of GAI in materials science will be discussed in Section 4. Finally, the significant conclusions from this review are presented in Section 5.

2. Various GAI methods and their applications to materials science

Nowadays, researchers modify the learning algorithms or model structures of GAI to fit specific domains to tackle specific tasks. Thanks to the accumulation of big data and the abundance of computing resources, large-scale models such as ChatGPT and GPT-4 are proposed to improve the generalization ability to solve various tasks, which can be regarded as the beginning of general GAI. In this section, the development process of task-specific and general GAI research is reviewed, especially the applications of GAI in materials research, as shown in Fig. 1.

2.1. Generative Adversarial Network

Generative Adversarial Networks (GANs) [18] were first introduced by Goodfellow *et al.* in 2014. As shown in Fig. 2(a), GAN is a deep learning architecture comprised of a generator and a discriminator. The generator creates synthetic data that mimics the real data, while the discriminator distinguishes between real and fake data. During the adversarial training process, the generator and discriminator engage in a competition, wherein the generator strives to produce more realistic data and the discriminator aims to accurately classify whether the data is real or fake. As a generative model, GAN is often used to generate samples for data augmentation and data pre-processing methods in deep learning, and has

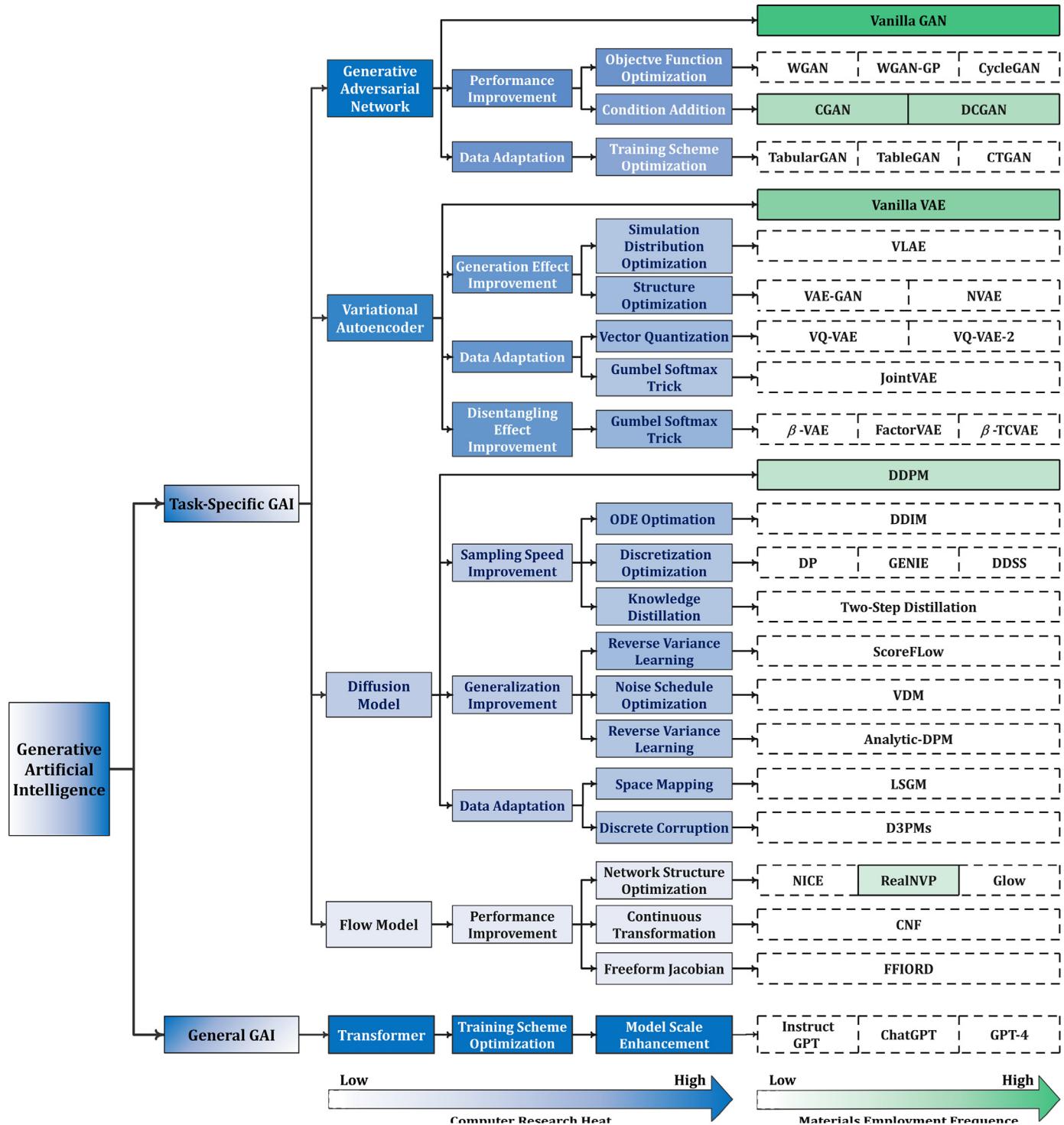
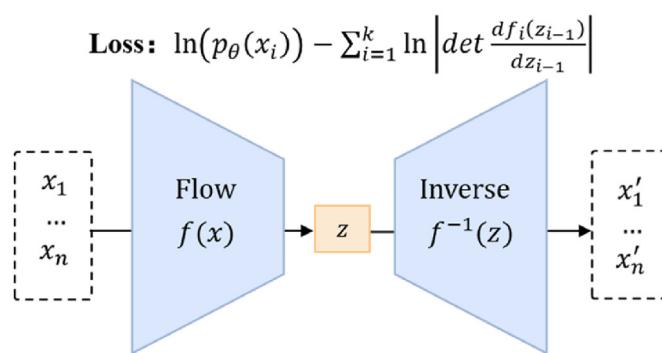
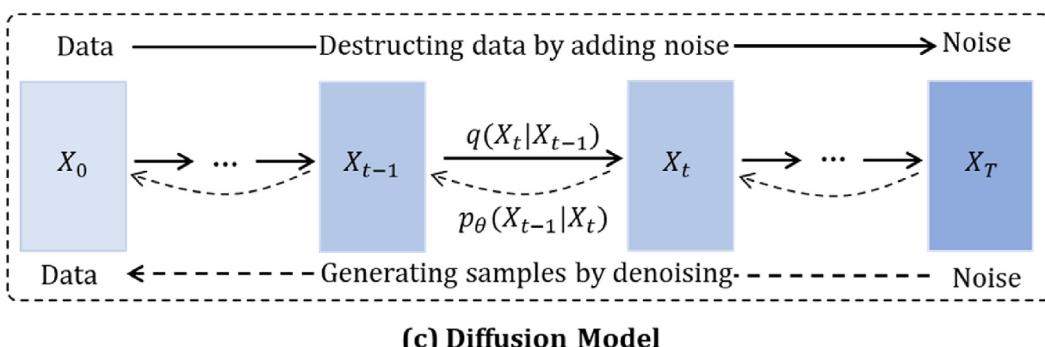
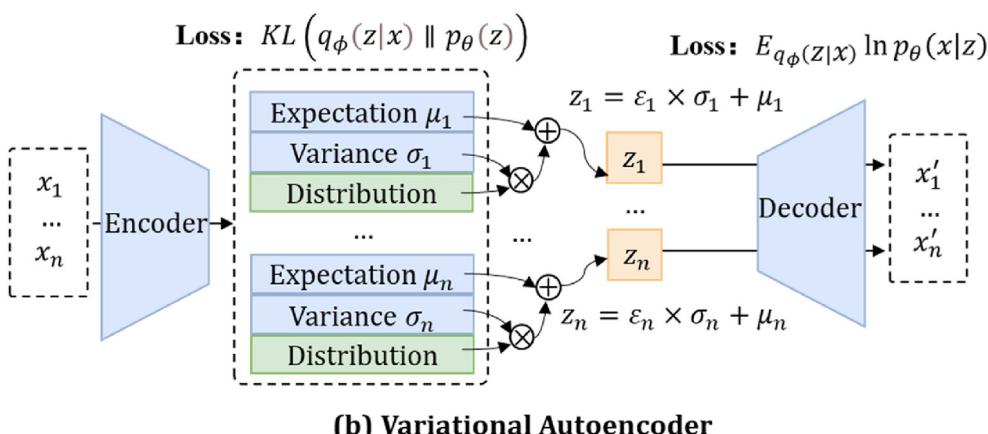
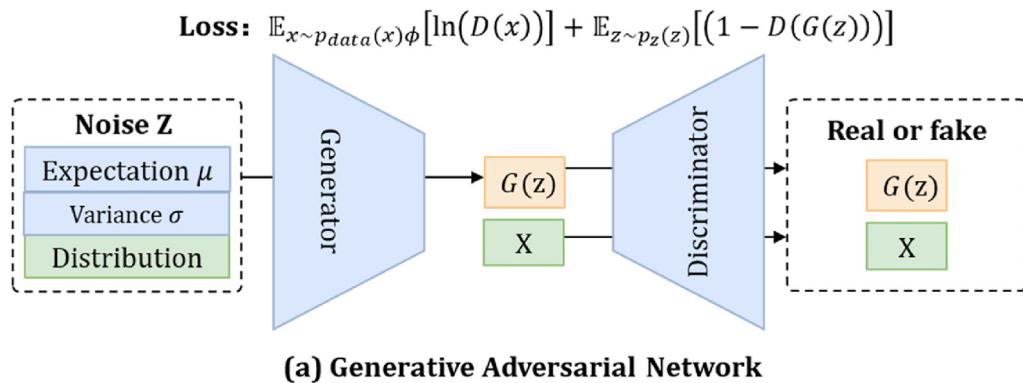


Fig. 1. Development of GAI and its application in materials science. The left arrow means the research heat in computer science. The right arrow means the employment frequency in materials science.

been widely used in image processing, biomedicine, and other fields. However, the Vanilla GAN is immature, it still has some limitations such as data mismatch, unstable training, and mode collapse. To address these issues, researchers have taken various approaches and conducted extensive analysis and research. For instance, researchers have developed conditional GAN (CGAN) [19] and deep convolutional GAN (DCGAN) [20] to improve model performance by introducing additional constraints, mitigating the

issue of an overly unconstrained training process. To address the issue of unstable training and mode collapse, others have proposed alternative GAN models, including Wasserstein GAN (WGAN) [21], WGAN with gradient penalty (WGAN-GP) [22], and CycleGAN [23]. These models optimize the loss function to improve the performance of GAN. In addition, to adapt to tabular data, some researchers proposed TableGAN [24], TabularGAN (TGAN) [25], and conditional tabular GAN (CTGAN) [26] to tackle generating different

**(d) Flow Model****Fig. 2.** Schematic diagram of vanilla generative models: (a) Generative adversarial network; (b) variational autoencoder; (c) diffusion model; (d) flow model.

types of data by optimizing training schemes. Fig. 3 lists the detailed comparison of GAN flavors.

Nowadays, GAN and its variants are employed to generate data and reverse design in the field of materials science. In data augmentation, GAN and TGAN are two popular tools for generating additional training data for performance prediction models. For example, Ma *et al.* [27] used GAN to expand 47 images of polycrystalline iron to 136, and performance of the model based on the generated data with 35 % of the real data reached the same level of performance as that based only on the real data, indicating that GAN have certain potential in generating images. Marani *et al.* [28] used TGAN to generate 6513 new data points based on 810 experimental values to predict the compressive strength of ultra-high performance concrete. The results demonstrated that the ML model trained on the generated data exhibited comparable prediction performance to that trained with real data, which shows that TGAN can accurately capture the distribution of real datasets and generate high-quality synthetic data. Yang *et al.* [29] proposed a two-step data augmentation method based on GAN to enhance the hardness prediction data of high-entropy alloys. The experimental results showed that the RMSE of the ML model constructed based on the original training set and generated data is 0.123, which is 6.1 % less accurate than the model constructed using only

real data. Furthermore, GAN have been used in materials reverse design to more efficiently sample materials space and generate hypothetical materials. For instance, Dan *et al.* [30] proposed a generative model MatGAN for data augmentation in the materials field. Based on this model, they sampled the materials databases ICSD, MP, and OQMD, and finally generated 2 million hypothetical inorganic materials. Subsequently, they explored the two-dimensional (2D) material space and used MatGAN to generate 2.65 million 2D material samples, ultimately discovering 26,489 new potential 2D materials, of which 1 485 2D materials had an area under the curve (AUC) score greater than 0.95 and 12 2D/layered materials were confirmed through DFT calculations [31]. These works demonstrate the potential of GAN in expanding materials reverse design space and screening large-scale inorganic materials datasets. Kim *et al.* [32] directly used GAN to predict ternary compounds containing only Mg, Mn, and O atoms, ultimately discovering 23 crystal structures suitable for photoelectrochemical anode applications. Long *et al.* [33] developed a reverse design framework CCDCGAN based on DCGAN for predicting crystal structures with target properties, and successfully generating crystal structures with low formation energy that had not been previously reported, thus accelerating the discovery of new phases.

Time	Model	Advantage	Disadvantage	Application
2014	Vallina GAN	Ability to generate high-definition images	Training unstably; Mode collapse	Materials Data Augmentation [30, 31]
Condition Addition	CGAN	Ability to generate specific types of samples using labeled data	High requirement for labeled data	Materials Reverse Design[32]
2016	DCGAN	Stable and efficient training process	Low quality for generated images	Materials Reverse Design[33]
Objective Function Optimization	WGAN	Stable training process	Lower diversity for generated samples	-
2018	WGAN-GP	Stable training process	Convergence speed slowly	-
Training Scheme Optimization	CycleGAN	Ability to perform image-to-image translation without paired data	Low resolution for generated images	-
	TableGAN	Adding a classifier on top of DCGAN to ensure the validity of generated data	Limited to tabular data	-
	TGAN	Generate tabular data with multinomial/discrete and continuous variables	Low quality for generated discrete data	-
	CTGAN	Handle mixed data types	Longer training time	-

Fig. 3. Development of GAN and its examples.

In summary, GAN has shown promising capability in image generation, tabular data augmentation, and new materials discovery and design. The trained GAN can better predict and optimize the materials properties, overcoming the limitations of small experimental datasets. New materials can be generated by learning the structure information from existing materials and further exploring the latent space of materials. However, challenges still exist in the application of GAN in materials science. Sometimes the new materials generated by GAN are difficult to synthesize and apply experimentally. In addition, due to the lack of reliable and consistent evaluation indicators, the prediction results are sometimes inaccurate. Thus, optimizing the network architecture, objective function, and proposing new improvement methods are necessary to advance the application of GAN in materials science.

2.2. Variational autoencoder

The Variational Autoencoder (VAE) [34] was first introduced by Diederik P. Kingma *et al.* in 2013. This model provides a concise way of capturing the essential low-dimensional information from the data, which can be used to generate new samples through simple manipulation of the learned low-dimensional representations via a decoder. VAE has emerged as a crucial method for both representation learning and generative modeling. As shown in Fig. 2(b), VAE is composed of an encoder and a decoder. The encoder extracts the mean and variance of the latent variables that determine its properties from the data through a neural network. The decoder adds Gaussian noise to the encoded information to generate new data. Unlike traditional autoencoders [35], which only compress the image to a point in the latent space, variational autoencoders force the latent variables of the latent space to a standard normal distribution. The encoder is no longer given a point but a distribution, allowing the model to learn a smooth latent state representation of the input data. This enables the model to cover unseen samples in the input data. However, the original VAE has limitations such as unsatisfactory generation effect, unsuitability for discrete data, and an insufficient disentangling effect. In order to solve these problems, researchers proposed various improvements, such as VAE-GAN [36], VLAЕ [37], NVAE [38], etc., which enhance the effect of model generation. Additionally, VQ-VAE [39], JointVAE [40], VQ-VAE-2 [41], and other approaches have been developed to adapt to discrete data. Lastly, a variety of methods have been suggested to improve the model's disentangling effect, including β -VAE [42], FactorVAE [43], and β -TCVAE [44]. Fig. 4 provides a detailed comparison of VAE and its related improved models.

In materials science, VAE has become a promising tool to solve material problems. The VAE-based models transform discrete material data representation into multi-dimensional continuous variables through an encoder, and then convert the multi-dimensional continuous variables into material representation through a decoder. Due to the distribution constraints imposed by the model on its latent space, the latent space becomes a data distribution rather than being discrete data points. Therefore, in terms of material generation, new material structures can be generated by performing simple operations on latent space distribution, to effectively explore the compound space, which was not originally covered. For example, Gómez-Bombarelli *et al.* [45] applied VAE to the design of drug-like molecules. By using an encoder network to convert a discrete molecular representation into a continuous vector in the latent space, and then performing simple operations on the latent continuous vector, such as perturbing known chemical structures, or interpolating between molecules. Then the modified vector can be converted back into a new discrete molecular representation through a decoder. Finally, a predictor module is used to predict the chemical properties of the new molecular

representation, thus effectively searching for candidate materials with higher target performance. Noh *et al.* [46] proposed a VAE strategy framework for the inverse design of solid-state materials, which efficiently explores the chemical compositional space by learning the distributions of known materials for crystal structure prediction. This method successfully discovered high-performance vanadium oxides that were previously identified in experiments and generated about 20 000 hypothetical materials, generating more than 40 new metastable high-performance vanadium oxides with energy above hull ≤ 80 meV/atom, i.e., potentially synthesizable. Additionally, due to the natural disentanglement phenomenon exhibited by the latent space of VAEs, Sardeshmukh *et al.* [47] designed a VAE model based on style loss to learn interpretable disentangled latent representations of material microstructures, capturing important attributes responsible for mechanical properties and generating structures with particular properties. This approach of using latent variables to explore the connection between the microstructure and properties of materials holds promise for devising reverse inference and is of great significance for designing materials and products with targeted performance. When representing compounds as vectors of components' proportions, the problem of discrete features and severe sparsity arises. The traditional VAE model used for processing continuous image data cannot directly generate discrete data. To address this, Oubari *et al.* [48] proposed Binded-VAE, a model specifically designed for generating highly sparse discrete datasets.

The prospects of using VAEs for material design are promising, as they have shown to be effective in generating novel materials with desired properties. However, there are still challenges that need to be addressed for VAEs to be fully effective in material design. Firstly, the VAE architecture needs to be optimized for materials data, which is different from image or text data commonly used in VAE applications. Secondly, the objective functions used to optimize the VAE should be carefully designed to accurately reflect the desired material properties. Thirdly, VAEs should be able to handle and generate data with high dimensionality, such as crystal structures, and effectively model the correlations between different properties of a material. Finally, it is important to validate the generated materials to ensure that they both have desired properties and are stable in real-world conditions. Despite these challenges, VAEs have the potential to revolutionize the way materials are designed and discovered with further advancements.

2.3. Diffusion model

The Denoising Diffusion Probabilistic Model (DDPM) [49], proposed by Jonathan Ho and Pieter Abbeel in 2020, outperformed GAN [50] in image synthesis upon its release. As a result, many subsequent studies on image generation have shifted towards DDPM-based research, leading to impressive results for the diffusion model in various generative modeling tasks. The diffusion model acquires the distribution of some crucial parameters through a neural network, as illustrated in Fig. 2 (c). During the training process, the model first performs a forward pass by gradually adding noise to the original data. At each step, the parameters are iterated through a Markov process until the data is transformed into a form that can be approximately represented by pure Gaussian noise. Following the forward process, the model then engages in a backward pass whereby the recently acquired representation form undergoes a gradual reduction of noise through decoding, resulting in the generation of new data. By aiming to achieve a greater resemblance between the new and original data, the model's parameters are continuously optimized, leading to an ongoing improvement in the quality of the generated content. The

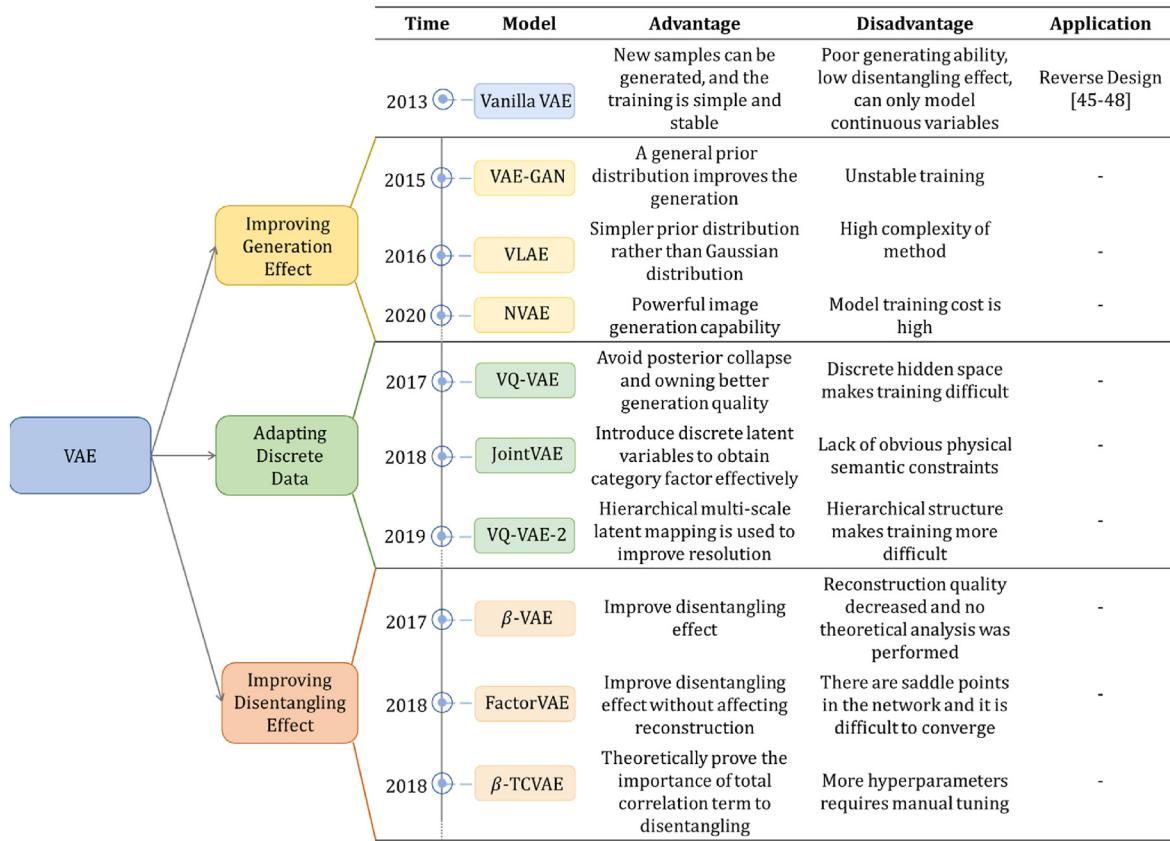


Fig. 4. The development of VAE and its application examples.

training process of diffusion models is known to provide a higher level of stability when compared to GANs, while its ability to generate a broad range of high-quality samples surpasses that of VAEs. However, the original diffusion model has three main shortcomings: slow sampling speed, poor maximum likelihood, and weak data generalization ability. Researchers have addressed these three shortcomings with various methods tailored to the task at hand. For instance, to improve sampling speed, methods such as DDIM [51], DP [52], DDSS [53], GENIE [54] and Two-step distillation [55] have been proposed. For enhancing data generalization, methods like LSGM [56] and D3PMs [57] have been developed. And Improved DDPM [58], VDM [59], ScoreFlow [60] and Analytic-DPM [61] have been put forward to strengthen the maximum likelihood. Fig. 5 shows the details of the diffusion model and its related modifications.

The diffusion model, with its unique network structure, requires only one forward and corresponding inverse process training to achieve stable training loss and excellent model performance. It can be combined with material characteristics to generate new materials that meet the desired targets. Lim *et al.* [62] used the diffusion model to generate an optimal microstructure with multi-functionality. By conducting multi-objective optimization, the light sensitivity and fracture toughness of the material were improved, offering a novel approach to characterize performance-based composite material microstructures. For proteins as macromolecular compounds, Anand *et al.* [63] introduced a fully data-driven diffusion model for protein structure, sequence, and rotamers, enabling the generation of highly realistic proteins across the full range of domains in the Protein DataBank. Schneuing *et al.* [64] proposed the DiffSBDD model that respects translation, rotation, reflection, and permutation equivariance, which is capable of

generating multiple ligands with high binding affinity for specific protein targets. Similarly, to design proteins of high fidelity as regards both sequence and structure, Shi *et al.* [65] proposed a new collaborative design method of protein sequence and structure based on diffusion model, PROTSEED, which can translate protein sequence and structure from random initialization to specific desired state based on given prior context features.

In summary, the diffusion model has garnered increasing attention and has been applied to material generation due to its outstanding performance. The diffusion model, after being adapted to material properties, can generate specific materials that meet the desired requirements, thereby exploring novel materials with improved properties. However, compared with other generative models, the diffusion model is still hindered by issues such as large memory consumption and computational costs, and is not yet mature enough in handling non-image data such as material data. Therefore, further optimization of the diffusion model's network architecture, objective functions, and data applicability is required to effectively tackle material-related challenges.

2.4. Flow model

Dinh Laurent *et al.* [66] introduced Non-Linear Independent Components Estimation (NICE) in 2015, which marked the birth of flow models. Unlike other generative models that attempt to approximate the true distribution of the data, flow models directly solve the mapping transformation between two distributions by manipulating the Jacobian determinant. This unique approach has made flow models a distinct and noteworthy type of generative model. Flow model establishes the mapping relationship between the distribution of data representation and its latent variable

Time	Model	Characteristic	Application
2020	DDPM	Easy training and can produce clear images.	Reverse design[63–65]
2021	Improved DDPM	A simple reparameterization and a hybrid learning objective are used to learn the reverse process variances.	Reverse design[62]
2021	VDM	Learnable diffusion specification, Fourier features for fine-scale modeling and architectural innovations are incorporated in the model.	-
2021	ScoreFlow	Introduce an efficient training objective termed likelihood weighting while lack theoretical proof.	-
2022	Analytic-DPM	A training-free inference framework which improves the log-likelihood, produces high-quality samples, and enjoys a $20\times$ to $80\times$ speed up.	-
2021	LSGM	Perform diffusion networks in hidden space, but currently only applies to image data.	-
2021	D3PMs	Introduce diffusion-like generative models for discrete data, but still inferior to strong autoregressive models.	-
2020	DDIM	Fewer sampling steps produce higher quality samples.	-
2021	DP	Introduce an exact dynamic programming algorithm to find the optimal discrete time schedules. But log-likelihoods and FID scores do not match.	-
2022	DDSS	Optimize fast samplers by differentiating through sample quality scores. But require extensive memory or computation requirements.	-
2022	GENIE	Second-order ODE solver for fast DDM sampling. But is still slower than approaches that abandon the differential equation framework of DDMs altogether.	-
2022	Tow-step distillation	Reduce the inference cost by at least an order of magnitude.	-

Fig. 5. The development of diffusion models and their applications.

distribution through a series of reversible transformations. Based on this, new content can be generated by sampling from the latent variable distribution. The structure of flow model is shown in Fig. 2 (d). Flow model has two characteristics. One is that it can calculate the distribution of latent variables that determine data representation, which ensures that the generated content and training data come from the same distribution. The second is that the entire model is built by a series of reversible changes and the data is converted to a priori distribution through a bijective function, which improves the interpretability of the model. Researchers have studied the performance improvement of flow models and proposed RealNVP [67] and Glow [68] by optimizing the network structure based on NICE. In addition, Chen *et al.* [69] combined differential equations and deep learning to construct continuous normalizing flows, thus avoiding the single-unit bottleneck problem in traditional normalizing flows. Grathwohl *et al.* [70] further constructed an unbiased estimator with $O(D)$ cost, which has been improved in high-dimensional density estimation, image generation and variational inference. Fig. 6 shows the comparison of classic Flow-based models. In contrast to other generative models, the flow model only uses a reversible encoder to complete the construction of the model, which makes the design of the model more delicate. However, this also raises the issue of more parameters and increased computing cost, which is the crucial challenge that still needs to be addressed.

Flow models are also useful in solving problems in the field of materials science. Material data have the issue of small data size,

which is an important factor to hinder the application of ML in materials science. Ohno [71] applied normalizing flows to the dataset of ionic conductivity of materials and augmented the data by using a real-valued non-volume preserving model (RealNVP) as the normalizing flow. This effectively solves the problem of insufficient training data in material performance prediction.

The utilization of flow models in materials science may provide a promising avenue driving advancements in the field. However, it is still limited by many aspects, such as model, computational, and material structure complexity and optimal parameter setting issues. In summary, the application of flow models in the field of material application requires comprehensive consideration of material characteristics, physical mechanisms, computational efficiency, and reliability, so as to realize its full potential in materials research.

2.5. General GAI

The advancement of big data and data representation technologies has led to the pursuit of generating human-understandable language from input data patterns and structures, with the capability to accomplish goals across diverse environments. This aim seeks to surpass the current language generation paradigm, which is limited to fitting sample distributions for specific tasks. Large language models (LLMs), with their powerful self-supervised learning ability, have demonstrated a remarkable capacity for “general” intelligence, enabling them to perform complex tasks and

Model	NICE	RealNVP	Glow	CNF	FFORD
Advantage	The encoder is reversible and directly fits the probability distribution	The introduction of convolution layer can better process the image and reduce the amount of computation	Generate HD images on complex data sets, enabling data interpolation and data modification	Avoid single-unit bottleneck problems, and it is easy to train	Continuous-time invertible with unbiased density estimation and one-pass sampling
Disadvantage	The model is simple, computationally intensive and valid only on simple data sets	The network structure is complex	The amount of calculation is still large and the training time is long	Only verified on simple data sets	The calculation speed is slow
Application	Material Data Augmentation [71]				

Fig. 6. Comparison of classic Flow-based models.

answer questions without intrinsic motivation or objectives. Beyond its mastery of language, LLMs represented by ChatGPT based on transformer can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting, and are seen as the prototype of general GAI. Nowadays, Google proposed their own LLM named Claude, which has impressive dialogue and task processing skills, to compete with ChatGPT or other advanced models proposed by OpenAI. Meanwhile, some Chinese companies such as HUAWEI, Baidu, Alibaba and Tencent all proposed their own LLMs, aiming to build industrial models and help industrialize AI large models. Here, we take ChatGPT as an example to discuss the development process of general GAI.

Fig. 7 illustrates the evolution of generic GAI. Initially, Generative Pre-trained Transformer (GPT) displayed its potential for task-specific natural language generation via unsupervised pre-training and fine-tuning on downstream tasks. It utilizes 12 transformer-decoder layers to execute next word prediction and coherent text generation. The purpose of fine-tuning is to adapt it to a specific task based on pre-training. GPT-2 [72] expands the model structure and parameters from its predecessor and trains on various datasets beyond web text. Despite exhibiting advanced results with zero-shot learning, it still falls under the category of task-specific GAI. The GPT-3 [73] leverages Prompt to minimize the need for large supervised datasets. It models the linguistic structure of text probability, enabling the probabilistic predictions. Specifically, the approach allows the language model to be pre-trained on a large amount of text, enabling it to perform few-shot or zero-shot learning and adapt to new scenarios quickly by defining a new cue template, even with limited or no labeled data available. Compared to “Pre-train + Fine-tune” [74], Prompt extensively taps into the prior knowledge of language model, allowing it to “Recall” the information learned during the pre-training stage when executing generation. Consequently, with appropriate prompts, the LLMs can handle numerous downstream tasks in a completely unsupervised manner. In materials science, Xie *et al.* [75] proposed a novel NLP task called structured information inference, which aims to address the complexities of information extraction at the device level in material science. Then, GPT-3 was leveraged and trained on perovskite solar cell FAIR dataset, which gained more

than 91 % accuracy. The results show that such LLMs can obtain acceptable performance to explore scientific knowledge and design novel materials. However, since it is hard to update the parameters of LLMs, such models fail to solve complicated materials tasks effectively. Moreover, the lack of domain knowledge aggravates the gap between model analysis results and real situation.

The advent of Reinforcement Learning from Human Feedback (RLHF) facilitates the shift from task-specific to general GAI in the case of InstructGPT [15]. RLHF optimizes the original problem by converting non-differentiable objectives in language generation task into sequential decision processes. Compared to traditional RL, RLHF is better able to align with human preferences and learn fluently from various types of rewards. Organizing human preferences on demand is integral to creating a true general-purpose intelligence. By using RLHF for fine-tuning, InstructGPT can make decisions that are consistent with human preferences, resulting in an intelligent system with greater adaptability, interpretability, and reliability. Furthermore, optimization of the fine-tuning process improves the exploration efficiency of the model, allowing it to achieve better performance but fewer parameters compared to GPT-3.

ChatGPT uses the same methods as InstructGPT but with slight difference in the data collection setup. As depicted in Fig. 7, an initial model is first trained *via* supervised fine-tuning: human AI trainers provide conversations in which they play both sides—the user and an AI assistant. This new dialogue dataset is mixed with the InstructGPT dataset and converted into a Q&A format. Then, reward models for reinforcement learning are created by collecting and ranking conversations between the AI trainer and the chatbot, along with alternative responses. Finally, the model is fine-tuned using proximal policy optimization (PPO) based on these reward models. Several iterations of the process resulted in a conversational generative model that performed with high efficacy. In addition to RLHF-based training techniques, the deployment of model is also crucial. ChatGPT adopts an iterative deployment approach to minimize the risks associated with generative models. This includes conducting a security assessment before deployment, executing beta testing with a small user group, and closely monitoring usage with retrospective reviews. For materials research, Hong *et al.* [76] explored some promising application of ChatGPT

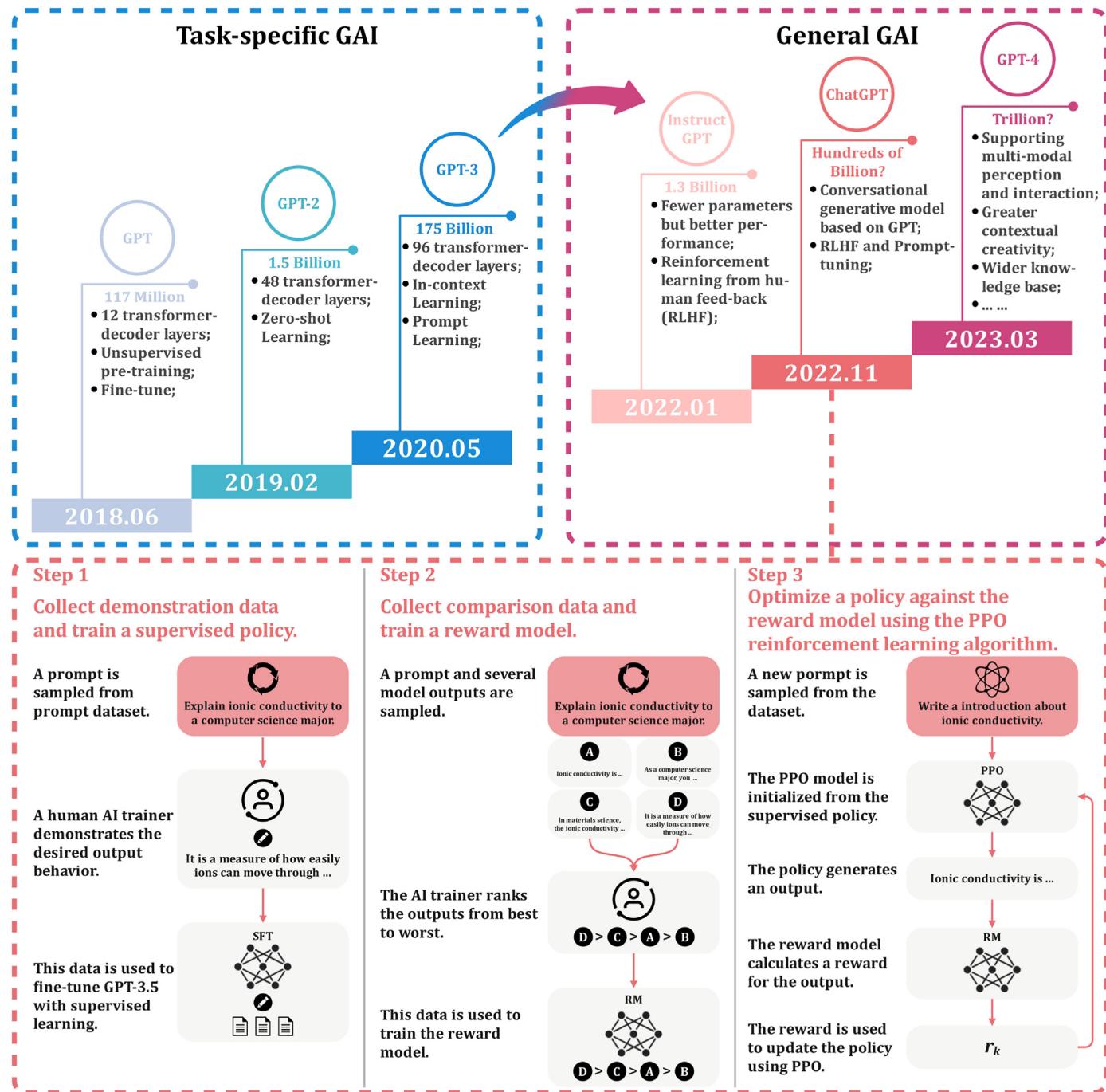


Fig. 7. Development of General GAI and the details of ChatGPT.

(namely building structures, writing codes for specific scientific software, and preparing data visualization scripts). The results show that such model may generate results with some errors, but it is still promising.

Recently, GPT-4, the latest model developed by OpenAI, was trained with an unprecedented scale of computations and data, surprisingly achieved human-like performance across almost all tasks and significantly outperformed its predecessors. The emergence of GPT-4 signifies a “general” advancement of general GAI based on the GPT series, encompassing the perception and generation of multimodal data such as text, images, and audio. The

formidable capabilities of GPT-4 in multimodal generation and conversational interactivity offer a promising outlook for materials science research. This breakthrough may drive the development of novel studies, such as crystal image analysis, crystallographic information file (CIF) generation, complex differential equation solving, and material Q&A systems, all of which could facilitate exploration of materials structure-activity relationships and discovery of new materials. Furthermore, ensuring the credibility, usability, and security of generated content remains a critical challenge that must be addressed for the future application of general GAs in materials science.

3. Future applications of general GAI in materials science: ChatGPT example

High-quality data is critical for materials science research, providing valuable insights for experiments and samples for ML. The advanced ability of general GAI to generate data, code, and model now offers materials researchers a great deal of convenience. In this section, we explore various applications of general GAI (taking ChatGPT as an example), namely generating materials data, solving differential equation, and querying FAQs, and analyze its reliability and availability for materials research and discovery.

3.1. Generating materials data

Materials research generates a large amount of data containing domain knowledge, which not only is the materials research foundation but also provide motivation for researchers. As a general GAI, ChatGPT has shown great capability of generating data, which has led to broad application prospects in materials science research. To this end, in this section ChatGPT is employed to generate CIFs and experimental data, and then evaluate its reliability and availability of materials data generation.

CIF is a file format proposed by the International Union for Crystallography (IUCr) to archive and distribute crystal structure information of materials. Such files include chemical formulas, unit cell parameters, space groups, atomic coordinates, position occupancies, etc., and have been widely employed to report crystal structures. Therefore, accurate CIF information is fundamental for scientific research on materials. Here, we utilize ChatGPT to generate CIF files for four systems of solid electrolytes *i.e.*, $\text{Li}_{0.5}\text{La}_{0.5}(\text{TiO}_3)$, $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$, $\text{Li}_{14}\text{ZnGe}_4\text{O}_{14}$ and $\text{NaZr}_2(\text{PO}_4)_3$, shown in Fig. 8. Although ChatGPT can generate the concept of full and partial occupancy of atom sites and the generated CIF files can be recognized by VESTA, their information is far from the required crystal structure, which is due to the lack of domain knowledge about crystal structures related to solid electrolytes.

To this end, we leverage Prompt to edit and improve the generation process, thus effectively modifying the generated results guided by the domain knowledge. Taking $\text{NaZr}_2(\text{PO}_4)_3$ as an example, we provide corresponding domain knowledge *e.g.*, rules related to space groups and key-value pairs, to modify the generated results, as follows:

Rule 1: The space group of $\text{NaZr}_2(\text{PO}_4)_3$ should be “ $\overline{R}\bar{3}C$ ”.

Rule 2: The migration ion Na occupies three different sites, namely Na1 sites (6b), Na2 sites (18e) and Na3 sites (36f), Zr ion occupies 12c site and the P ion occupies the 18e site.

Rule 3: The crystal structure should be a three-dimensional skeleton that forms a crystal structure through the vertices-sharing of ZrO_6 octahedra and PO_4 tetrahedra. The sodium ions are distributed in different interstices of the skeleton, forming Na (1) O_6 polyhedron, Na (2) O_8 polyhedron and Na (3) O_5 polyhedron.

Prompted by above rules, ChatGPT gradually corrected the CIF file. As shown in Fig. 8, **Rule 1** can correct the crystal structure of $\text{NaZr}_2(\text{PO}_4)_3$ effectively. This emphasizes that spatial symmetry reflected by the space group is critical for the crystal structure. Moreover, owing to **Rule 2**, ChatGPT can effectively modify the sites of skeleton atoms. After modifications following **Rule 1 and 2**, the crystal structure has better approximated the target structure. Compared to the unconstrained bond generation, **Rule 3** effectively prevented erroneous bonds. As the number of questions and the application of domain knowledge increased, the CIF files

generated by ChatGPT were progressively corrected and improved, particularly in terms of their spatial symmetry. As ChatGPT lacks domain knowledge about atomic arrangement, atomic number, and atomic bonding, even after multiple adjustments and questions, it only reduces the generation of some keys. Though the results above suggest that such general GAI can provide feedback to the user during the interaction process and has self-correction function, currently ChatGPT fails to generate compliant CIF for aiding in materials research and discovery.

Experimental data is the fundamental demand for materials research, but as experiments and measurements are time-consuming and labor-intensive, the materials sample size is commonly small. To this end, the powerful data generation capability of ChatGPT appears to offer a solution. By extracting data from published literature, ChatGPT can generate Excel and CSV files in the corresponding format. However, the accuracy of the generated data sets is not satisfactory. Despite outputting ionic conductivity and activation energy values based on the search criteria, they do not match the literature content. Furthermore, the reference titles and provenance information are often incorrect.

In summary, at present ChatGPT often outputs incorrect data on request, but can be used to quickly retrieve the original corresponding documents, which can save researchers time in data collection. With the introduction of GPT-4, it becomes possible to access to more types of data, yet the accuracy of the data also needs more experimental verification. It is believed that with the development of general GAI, GAI models can be employed to output higher-quality multi-modal data for researchers.

3.2. Solving differential equation

In materials science, mathematical equations are constructed to express numerical relationships between various attributes and performance parameters. Differential equations are widely used in material science. For example, Chen *et al.* [77] used Neural ordinary differential equations (ODEs) to simulate behavior in spintronic devices with high precision and time efficiency. Li *et al.* [78] investigated the effect of diaphragm thickness and diaphragm surface coating on dendrite growth using a phase field model with a non-linear partial differential equation (PDE). In this process of construction of a suitable differential equation, domain knowledge plays a significant role.

Naturally, it is crucial to solve these equations so that they approximate the true values. Nowadays, mainstream simulation software, such as ANSYS, Hyperworks, and COMSOL, have achieved high levels of success in simulation computing. ChatGPT also owns the powerful capability of mathematics, which can greatly reduce computational costs while providing accurate results. To this end, here we present validation experiments using ChatGPT to solve differential equations as an example, to explore the potential of this capability in materials science.

As illustrated in Fig. 9, after inputting the equation and its relevant parameters in the dialogue box, ChatGPT performs the necessary calculation steps and analysis. For simple differential equations, conversion to a computer-comprehensible language is not necessary, and ChatGPT can automatically convert and solve the equations. While evaluating the discrete error of the differential equation, due to computing power constraints, ChatGPT cannot provide a complete solution and calculation process. However, after entering the command “Generate a visualization file of the discrete errors,” ChatGPT generates the corresponding Python code that can produce the required result after execution. It is also possible to request an alternative approach to the problem for better outcomes. In the case of more complex Partial Differential Equations (PDEs), ChatGPT fails to produce the full solution process, indicating

Test

Sample Type	Li ₁₄ ZnGe ₄ O ₁₄	Li ₇ La ₃ Zr ₂ O ₁₂	Li _{0.5} La _{0.5} (TiO ₃)	NaZr ₂ (PO ₄) ₃
Generated Sample				
Target Sample				

Prompt

Interaction process

Can you generate a CIF file for **NaZr₂(PO₄)₃**?

Certainly! Here's the CIF file for NaZr₂(PO₄)₃:

Reply result

Rule1: Please change the space group to 'R -3 c'.

Sure! Here's the CIF file for NaZr₂(PO₄)₃ with the space group changed to R -3 c.

Reply result

Rule2: Please note that the migration ion Na occupies three different sites, namely Na1 site (6b), Na2 site (18e) and Na3 site (36f), Zr ion occupies 12c site and P ion occupies 18e site.

Thank you for letting me know. Here's the updated CIF file for NaZr₂(PO₄)₃ with the new site occupancies:

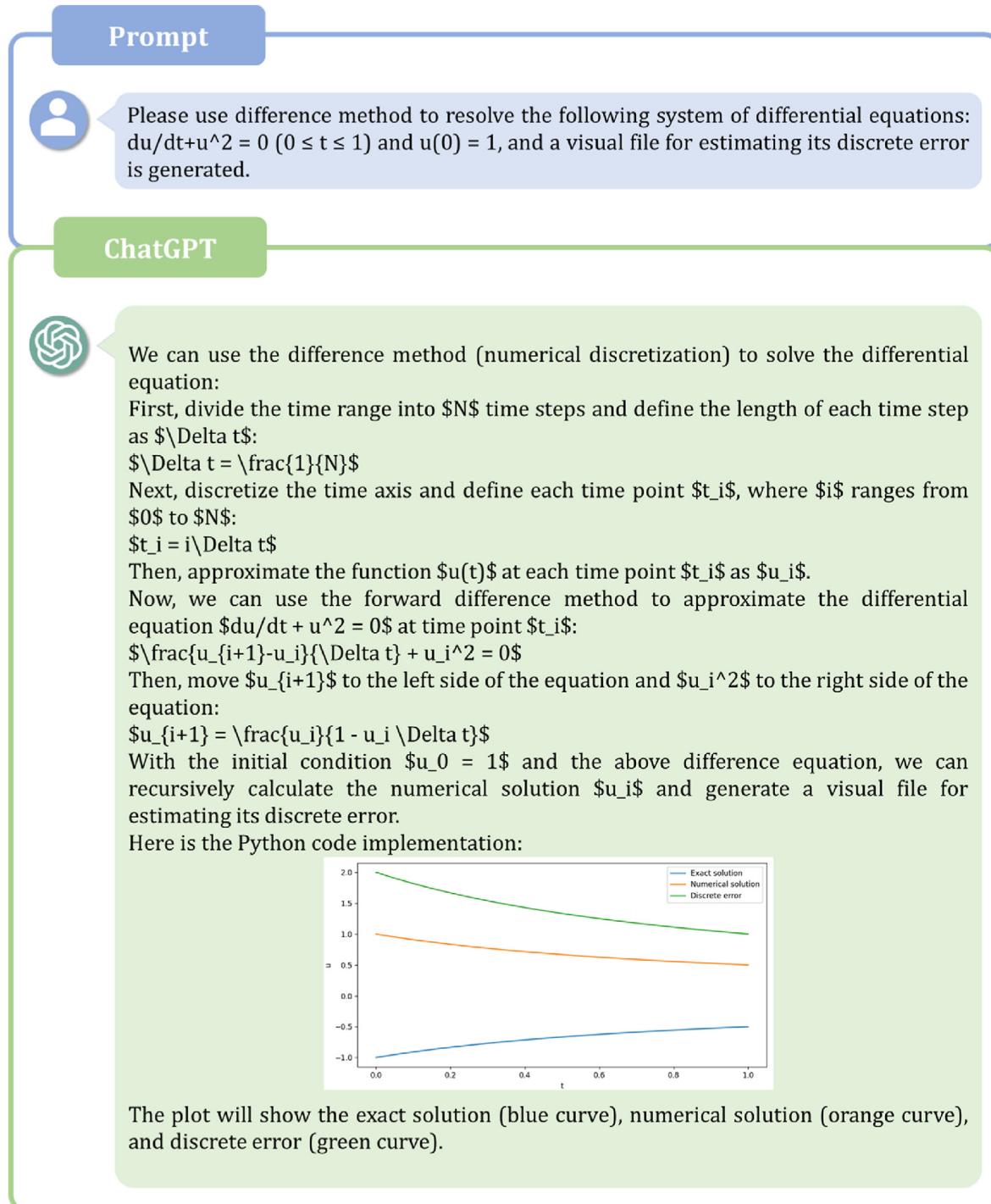
Reply result

Rule3: Please note that the crystal structure should be a three-dimensional skeleton of the crystal structure formed by the connection of ZrO₆ octahedron and PO₄ tetrahedron through vertex angles, and sodium ions are distributed in different gaps of the skeleton to form Na(1)O₆ polyhedron, Na(2)O₈ polyhedron and Na(3)O₅ polyhedron.

Thank you for the additional information. Here's the updated CIF file for NaZr₂(PO₄)₃ with the new structure description:

Reply result

Fig. 8. Comparison of generative and true CIF files and the revising process of NaZr₂(PO₄)₃.

**Fig. 9.** Process of solving differential equations by ChatGPT.

its low capability to PDE solution.

Since many areas of materials science involve differential equations involving physics and chemistry knowledge, at present the ChatGPT cannot directly support materials computations or simulations, but its capability to understand and generate limited computational content and programming, e.g., Python, code, which can be used with external software, is evident.

3.3. Querying materials FAQs

Modification of materials according to application requirements is one of the most important tasks in materials science. Thereinto, one of the crucial steps in material modification research is to analyze and employ publicly available modification strategies, which is the foundation of experiment design. However, this task can be challenging and time-consuming as it requires researchers to read numerous papers, reviews, and patents. Furthermore, important publications may be missing, e.g., due to limited access to

journals, which can result in suboptimal experimental designs or even duplication of completed work. Therefore, if ChatGPT could be used to query materials-science-related FAQs (Frequently Asked Questions), such as the modification strategy of the material, it would alleviate the effort to review the literature for researchers and thus facilitate experiment design.

The perovskite-type solid electrolyte $\text{Li}_{0.5}\text{La}_{0.5}(\text{TiO}_3)$ (LLTO) possesses stable physical and chemical properties, making it a promising candidate for the next generation of all-solid-state lithium-ion batteries. However, the ionic conductivity of LLTO falls short of meeting the requirements for high charge/discharge rates of lithium-ion batteries. Therefore, it is necessary to enhance the ionic conductivity of LLTO. Herein, we take LLTO solid electrolyte as an example, to validate the feasibility of querying material modification strategy through ChatGPT. As shown in Fig. 10 (left), ChatGPT collects LLTO's modification schemes from various perspectives, which provides reference for researchers. However, ChatGPT fails to offer a practical solution by considering the advantages and disadvantages of each scheme and cannot provide specific methods to determine whether we should increase or decrease the temperature, prolong, or shorten the annealing time. Moreover, ChatGPT responds that the ionic conductivity of LLTO can be improved by controlling the grain size and introducing a secondary phase but not how that can be achieved. When queried how to improve the ionic conductivity of $\text{Li}_{0.5}\text{La}_{0.5}(\text{TiO}_3)$ through doping, ChatGPT not only provided several doping schemes, but also described the mechanism of each doped ion, as shown in the right part of Fig. 10. However, ChatGPT failed to provide advantages and

disadvantages of various dopants.

The results show that ChatGPT has the potential to serve as a “document steward” for materials researchers, which can efficiently gather relevant information from numerous literature sources. Nevertheless, its current limitations include the inability to provide insight into the value of retrieved information and the difficulty in collecting important relevant details. This leads to having to continuously ask follow-up questions. Therefore, improvements in these areas through iterative upgrades can significantly enhance the utility of such general GAI models for materials researchers.

The current applications of ChatGPT in materials science demonstrate its rich knowledge base, making it suitable for providing users with basic knowledge of materials science. The Prompt paradigm enables users to obtain corresponding knowledge and correct misconceptions, while its data processing and code writing capabilities provide researchers with great convenience. However, its limitations are also apparent, i.e., it lacks specialized material domain knowledge and most of the answers fail to meet the needs of material researchers. Additionally, the availability of material text data generated by ChatGPT is limited and often contrary to material domain knowledge. The ability of ChatGPT to process complex data is also constrained by input format and computing power, often leading to incomplete output of calculation results. Despite these challenges, ChatGPT has demonstrated impressive potential for knowledge output and data calculation in the field of materials. Its future development is highly anticipated.

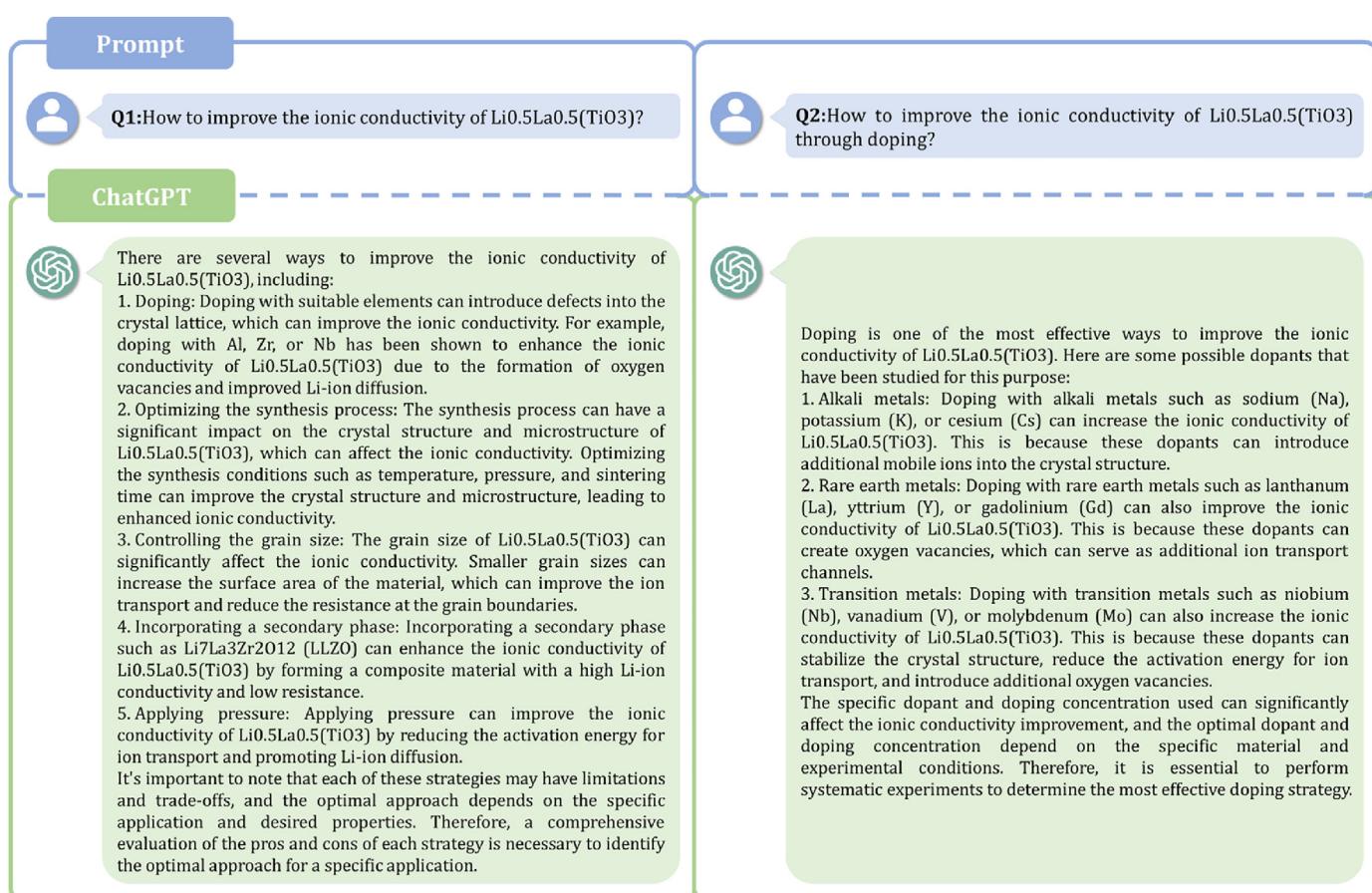


Fig. 10. Results of FAQs for ChatGPT. (Left) Question about How to improve the ionic conductivity of $\text{Li}_{0.5}\text{La}_{0.5}(\text{TiO}_3)$; (Right) Question about How to improve the ionic conductivity of $\text{Li}_{0.5}\text{La}_{0.5}(\text{TiO}_3)$ via doping.

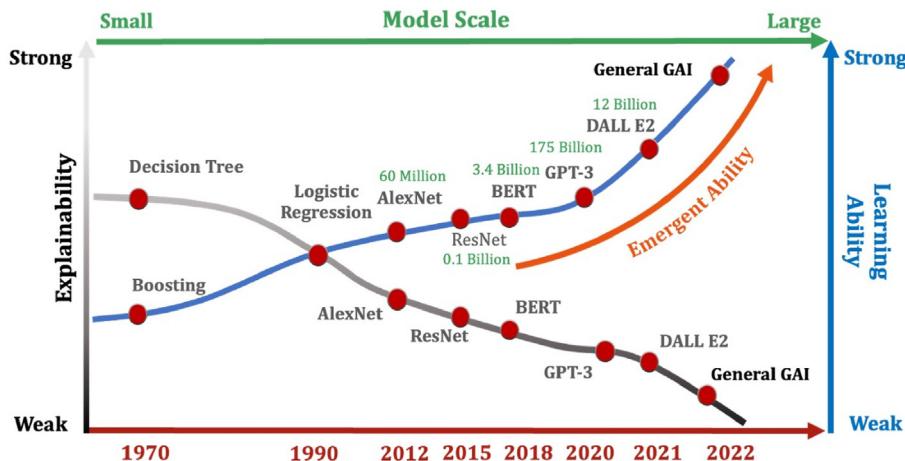


Fig. 11. Contradiction between model learning ability and interpretability.

4. Analysis of problems of GAI in materials science

GAI has participated in materials research and discovery widely. The feasibility of embedding domain knowledge (e.g., fundamental physical laws) into GAI models brings new possibilities for AI4Science to drive the research and development of novel materials. The experiments described above demonstrate that general GAI, such as ChatGPT, seem to be able to solve materials science problems from multiple aspects, which illustrates the potential to change the paradigm of materials research and development.

In the past, with the increase of model scale, the interpretability of the model greatly decreased, and its characteristics of “black box” profoundly affected the reliability, controllability, and credibility, as shown in Fig. 11. In contrast, the most recent developments in GAI revealed new behavior emerging with further model scale increase. The latest models, such as GPT, demonstrate extraordinary ability in data representation, data analysis and few-shot even zero-shot learning, which never occurs in smaller models [79]. This induces that people begin to consider whether and where GAI should be utilized. To this end, Liu *et al.* [2,80] proposed the framework of machine learning embedded with materials domain knowledge, which aims to transfer “black box” to “gray box” or even “white box” via embedding various domain knowledge into the entire process of ML. In this way, three key issues in materials science can be effectively reconciled and the interpretability, credibility and robustness of ML can be greatly improved. In 2017, we have pointed out several common problems of ML in materials science [1], which is still applicable to current GAI in materials science. Following this, we here discuss and analyze the problems of the future development of GAI and its applications in materials science, then provide possible solutions.

4.1. High-quality data and domain knowledge

Data representation affects the amount of valid information in samples, which in turn determines the data quality [81,82]. Efficient data representation for unstructured materials data facilitates GAI to fit the characteristics of the original samples more accurately through Gaussian distribution. Note that the paradigm from “Fitting-Generation” is appropriate for the samples with strong correlation between features (e.g., image, video). However, the high complexity of the structure-activity relationship latent in materials data makes it difficult to accurately fit the association between descriptors (features) through Gaussian distribution. Hence, the

generated samples leveraging the low-quality samples will amplify the defect of the original samples and result in serious contradiction with the original samples or domain knowledge. To this end, domain knowledge can be symbolized and integrated into data representation methods, which enables GAI models to learn basic physical laws while retaining the effective information and eliminating redundant information. Moreover, the data quality detection method with embedded domain knowledge [83] can effectively improve the data quality via identifying possible abnormal data in samples and remove abnormal samples under the guidance of materials domain knowledge. Through these approaches, GAI can generate new samples following domain knowledge, and thus build a virtuous circle that original and generated samples can be merged.

On the other hand, the objectivity, authenticity, and representation pattern of domain knowledge are key issues for the application of GAI in materials science. Embedding wrong domain knowledge into GAI models will hinder them from mining the structure-activity relationships latent in the materials data. Domain knowledge is mostly contained in scientific literature. Although researchers have employed large-scale language models to mine historical data or empirical knowledge from materials scientific literature [84–87], insufficient attention is often paid to the quality of domain knowledge. Therefore, it will be helpful for the improvement of the domain knowledge quality to quantify the credibility of materials scientific literature and perform the credibility analysis, which can be conducive to researchers to screen out domain knowledge with higher credibility from massive literature according to task requirements [88].

4.2. Generalization ability

Generalization ability is a vital metric for measuring the performance of AI models. The generalization ability of GAI models is usually expressed as the distribution of generated samples approximating the inputs with some extrapolation capability. When a GAI model is successful, it should interpolate or even extrapolate beyond the original training data which will be a continuation of the original data distribution. The generated materials data should have a similar or the same structure as the original and it should be the interpolation or extrapolation of inputs. For example, Ref. [33] developed a GAN-based inverse design framework to generate distinct crystal structures and applied it to the binary Bi-Se system. Ref. [89] proposed an approach combining

computational chemistry and GAN to generate new catalytic surfaces in an extrapolative manner. It is worth noting that the discrepancy in the generalization ability between the GAI models focusing on specific materials and general GAI models. The former have poor generalization ability to other types of materials yet generate highly accurate content for the target materials while the latter have strong generalization ability but with questionable reliability of the generated content. Moreover, how to effectively evaluate the generalization ability of the GAI models is an urgent problem to be handled. Therefore, it is necessary to comprehensively consider the sample quality, sample content, and GAI models. Embedding domain knowledge and common-sense constraints into GAI should also be explored, as that would make outputs more accurate, diverse, and controllable.

4.3. Interpretability and credibility

Although GAI models have the potential to revolutionize the current paradigm of materials research and discovery, there is an ongoing concern about the credibility and interpretability of GAI models. Model interpretability refers to the ability of a model to interpret the correlation between the input data and the target, as well as the semantic representation of the feature space and the behavior of the model during the learning process, which is widely discussed as a key factor of credibility. This enables the model builder or user to understand the learned features, which in turn can inform decisions about the model's performance and how to improve it. Therefore, how to explain and improve the interpretability of generative models is an urgent problem that needs to be addressed. Researchers [42,43,90,91] have attempted to focus on the semantic features contained in the learned latent space and manipulating them to explain the internal feature processing of the model. Note that the GAI model has extremely high latent space dimension and complex internal semantics, so how to interpret its internal semantic features is a key problem. Therefore, in addition to understanding the internal mechanism of the GAI models, it is necessary to examine their generation process and understand how it obtains the results step by step, e.g., a dynamic evaluation mechanism can be introduced to improve the interpretability, which will enable continuous interaction with the model and provide self-explanatory outputs. Moreover, embedding domain knowledge [92] into GAI models can also be regarded as an effective way. By these means, researchers can gain more accurate and reliable insights from their outputs, facilitating the development of new materials. Through such models we can also comprehend which features (descriptors) are crucial for the model.

4.4. Usability

Usability measures the difficulty of employing GAI models to deal with practical problems, which manifests in two aspects. 1) Complex and unpredictable generation process requires to carefully screen out reliable results from numerous candidate generated data. For example, Ref. [31] leveraged GAN to generate 267 489 new potential compositions while only 92 novel materials were confirmed by DFT formation energy calculation. Ref. [93] utilized GAN to generate 10 million samples, where appropriate 4.1 million samples can be obtained for downstream screening. Finally, only 506 materials were verified by phonon dispersion calculation. 2) Though deeper GAI models have significant ability to data representation and exploration of latent structure-activity relationships, the optimization space of hyperparameters and the demand of labeled data can experience an exponential surge. Therefore, in

order to improve the usability of GAI models, it is necessary to introduce Neural Architecture Search (NAS) to achieve the best possible performance using limited computing resources in an automated manner with minimal manual intervention [94,95]. Moreover, the complexity of the preparation and data labeling will reduce the usability of the GAI model. To this end, besides embedding physical information into generative process for manipulating generated samples, active learning is an effective approach to the selection of generated unlabeled samples [96], which not only can improve the performance of candidate samples and reduce the number of candidates enormously but also reduce the demand of human cost via screening out high-quality unlabeled samples.

4.5. Resource cost

While GAI models have powerful generation ability, its resource consumption is a noteworthy problem, which includes training costs and operating costs. Thereinto, the training cost refers to the cost required for the model to produce a certain generation capacity, e.g., computing power requirements, equipment requirements, and training time costs. The operating cost of the model mainly considers the energy consumption required for the model operation. The advanced generation ability of GAI models is due to the utilization of high computing power, which enables such models to process vast amounts of data and learn complicated patterns. Therefore, improving the learning efficiency in a reasonable and effective manner is key to reducing these costs. For example, Ref. [97] proposed a standard pruning technique to compress and optimize the GAI models without sacrificing generation quality. Ref. [98] explored the relationship between optimizing the training efficiency and resource consumption of GAI models. Furthermore, the accumulation of large amount of labeled data requires significant time and human cost. Hence, it is important to deal with computing resources and the efficiency of model output also from the perspective of model compression and usage costs, to facilitate GAI to be highly integrated with different fields and produce new research paradigms in various fields in the future.

4.6. Security

GAI models mostly own the characteristics of “black box”, which raises security concerns [99]. For example, attackers can exploit generative models to create convincing fake text [98] (e.g., false news articles, social media posts, and phishing schemes), thus deceiving people into engaging in unlawful activities. Moreover, GAI models are vulnerable to attack with respect to data privacy and intellectual property, such as training data or model parameters. As a result, it is essential to conduct in-depth research on security of generative models and to strengthen defense measures against potential attacks. For example, Ref. [100] proposed defense techniques to safeguard GAI models, considering a trade-off between the utility and security of generative models. For materials science, the inputs of GAI models are mainly obtained from experiments or computation, for which it is necessary to consider the intellectual property rights and security. In the future, the development of GAI will involve more content generation, including intellectual property, therefore, how to enhance the security of GAI models and safeguard both the generated content and internal data of the models is an urgent issue.

5. Conclusion

The appearance of large-scale generative models driven by Prompt paradigm and RLHF algorithm shifts the generation paradigm from “Fitting-Generation” to “Pretraining-Prompting-Generation”. Such models offer the hope of general GAI realization, that can facilitate the deep fusion between AI4Science and materials science and accelerate innovation in materials science research. Through reviewing the status of GAI, it can be seen that in computer science researchers improve the multi-aspect abilities of GAI models according to specific demands through different approaches (e.g., objective function optimization, condition addiction, etc.) and numerous improved models are proposed. The pros and cons of various generative models are summarized, aiming to offer a guidance to their applications in materials science. Following that, we review the research in GAI for materials science and find that materials researchers prefer to employ basic GAI models, while there are many promising improved and advanced GAI methods, which have not been explored yet. Specifically, we discuss the prospects of general GAI in materials science and leverage ChatGPT as an example for exploring whether it can solve the key issues in materials science (e.g., novel materials generation, solving differential equations, and querying FAQs). The results show that ChatGPT can effectively deal with materials problems and modify the generated results via Prompt to some extent. However, because of the lack of specific domain knowledge in its training dataset, ChatGPT currently still cannot be employed to handle the complicated or domain-knowledge-required materials problems. Finally, the six challenges of the use of GAI in materials science are discussed, i.e., the requirements of high-quality data and domain knowledge, the improvement of model generalization ability, interpretability and credibility, usability, security, and resource cost. We believe that embedding domain knowledge into GAI models would resolve many of these problems. The improvement of the data and knowledge quality can provide solid foundation for materials research and development. AI experts, materials domain experts, and non-experts will need to combine effort to resolve the identified issues. Overall, a general GAI is bound to promote the rapid development of AI4Science, and then open a new era of materials science research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported in part by National Natural Science Foundation of China [grant number 92270124, 52073169] and National Key Research and Development Program of China [grant number 2021YFB3802101] the Key Research Project of Zhejiang Laboratory [grant number 2021PEOAC02]. We appreciated the High-Performance Computing Center of Shanghai University and Shanghai Engineering Research Center of Intelligent Computing System for providing the computing resources and technical support.

References

- [1] Liu Yue, Zhao Tianlu, Ju Wangwei, Shi Siqi. Materials discovery and design using machine learning. *J. Materom.* 2017;3:159–77.
- [2] Liu Yue, Guo Biru, Zou Xinxin, Li Yajie, Shi Siqi. Machine learning assisted materials design and discovery for rechargeable batteries. *Energy Storage Mater* 2020;31:434–50.
- [3] Jovanović Mladen, Campbell Mark. Generative artificial intelligence: trends and prospects. *Computer* 2022;55:107–12.
- [4] Cao Hanqun, Tan Cheng, Gao Zhangyang, Chen Guangyong, Ann Heng Pheng, Stan Z Li. A survey on generative diffusion model. *arXiv:220902646* 2022.
- [5] Jabbar Abdul, Li Xi, Omar Bourahla. A survey on generative adversarial networks: variants, applications, and training. *ACM Comput Surv* 2021;54:1–49.
- [6] Sanchez-Lengeling Benjamin, Aspuru-Guzik Alán. Inverse molecular design using machine learning: generative models for matter engineering. *Science* 2018;361:360–5.
- [7] Luo Shitong, Shi Chence, Xu Minkai, Tang Jian. Predicting molecular conformation via dynamic graph score matching. *Adv Neural Inf Process Syst* 2021;34:19784–95.
- [8] Chen Litao, Zhang Wentao, Nie Zhiwei, Li Shunning, Pan Feng. Generative models for inverse design of inorganic solid materials. *J. Mater. Inform.* 2021;1:4.
- [9] Hoogeboom Emiel, Garcia Satorras Victor, Vignac Clément, Welling Max. Equivariant diffusion for molecule generation in 3D. In: Proceedings of the 39th international conference on machine learning, 162; 2022. p. 8867–87.
- [10] Zhao Yong, Edirisuriya M, Siriwardane Dilanga, Wu Zhenyao, Fu Nihang, Al-Fahdi Mohammed, Hu Ming, et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints. *npj Comput Mater* 2023;9:38.
- [11] Rick Stevens, Taylor Valerie, Nichols Jeff, Arthur Barney Maccabe, Katherine Yelick, Brown David. AI for science: report on the department of energy (DOE) town Halls on artificial intelligence (AI) for science. 2020. United States.
- [12] Liu Pengfei, Yuan Weizhe, Fu Jinlan, Jiang Zhengbao, Hayashi Hiroaki, Graham Neubig. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Comput Surv* 2023;55:1–35.
- [13] Wang Yaqing, Yao Quanming, James T. Kwok. Generalizing from a few examples: a survey on few-shot learning. *ACM Comput Surv* 2020;53(3):1–34.
- [14] Xian Yongqin, Christoph H. Lampert, Bernt Schiele, and Zeynep Akata. Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly. *IEEE Trans Pattern Anal Mach Intell* 2018;41:2251–65.
- [15] Ouyang Long, Wu Jeffrey, Jiang Xu, Almeida Diogo, Wainwright Carroll, Mishkin Pamela, et al. Training language models to follow instructions with human feedback. *Adv Neural Inf Process Syst* 2022;35:27730–44.
- [16] Bubeck Sébastien, Chandrasekaran Varun, Ronen Eldan, Gehre Johannes, Horvitz Eric, Kamar Ece, et al. Sparks of artificial general intelligence: early experiments with GPT-4. *arXiv:230312712* 2023.
- [17] Chengwei Qin, Aston Zhang, Zhuosheng Zhang, Jiaao Chen, Michihiro Yasunaga, and Difyi Yang. Is chatgpt a general-purpose natural language processing task solver?. *arXiv:230206476*.
- [18] Goodfellow Ian, Pouget-Abadie Jean, Mirza Mehdi, Xu Bing, Warde-Farley David, Ozair Sherjil, et al. Generative adversarial networks. *Commun ACM* 2020;63:139–44.
- [19] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv:14111784*.
- [20] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv:151106434*.
- [21] Martin Arjovsky, Chintala Soumith, Bottou Léon. Wasserstein generative adversarial networks. In: Proceedings of the 34th international conference on machine learning, 70; 2017. p. 214–23.
- [22] Gulrajani Ishaan, Ahmed Faruk, Martin Arjovsky, Vincent Dumoulin, Aaron C Courville. Improved training of wasserstein GANs. *Adv Neural Inf Process Syst* 2017;30.
- [23] Zhu Junyan, Park Taesung, Isola Phillip, Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. ICCV; 2017. p. 2223–32.
- [24] Noseong Park, Mahmoud Mohammadi, Kshitij Gorde, Sushil Jajodia, Hongkyu Park, and Youngmin Kim. Data synthesis based on generative adversarial networks. *arXiv:180603384*.
- [25] Lei Xu and Kalyan Veeramachaneni. Synthesizing tabular data using generative adversarial networks. *arXiv:181111264*.
- [26] Xu Lei, Skoularidou Maria, Cuesta-Infante Alfredo, Veeramachaneni Kalyan. Modeling tabular data using conditional GAN. *Adv Neural Inf Process Syst* 2019;32.
- [27] Ma Boyuan, Wei Xiaoyan, Liu Chun, Ban Xiaojuan, Huang Haiyou, Wang Hao, et al. Data augmentation in microscopic images for material data mining. *npj*

- Comput Mater* 2020;6:1–9.
- [28] Marani Afshin, Jamali Armin, Moncef L, Nehdi. Predicting ultra-high-performance concrete compressive strength using tabular generative adversarial networks. *Materials* 2020;13:4757.
- [29] Yang Zhiyuan, Li Shu, Li Shuai, Jia Yang, Liu Dongrong. A two-step data augmentation method based on generative adversarial network for hardness prediction of high entropy alloy. *Comput Mater Sci* 2023;220:112064.
- [30] Dan Yabo, Zhao Yong, Xiang Li, Li Shaobo, Hu Ming, Hu Jianjun. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. *npj Comput Mater* 2020;6:84.
- [31] Song Yuqi, Edirisuriya M, Siriwardane Dilanga, Zhao Yong, Hu Jianjun. Computational discovery of new 2D materials using deep learning generative models. *ACS Appl Mater Interfaces* 2021;13:53303–13.
- [32] Kim Sungwon, Noh Juwan, Gu Geun Ho, Aspuru-Guzik Alan, Jung Younsung. Generative adversarial networks for crystal structure prediction. *ACS Cent Sci* 2020;6:1412–20.
- [33] Long T, Fortunato NM, Opareh I, et al. Constrained crystals deep convolutional generative adversarial network for the inverse design of crystal structures. *npj Comput Mater* 2021;7:66.
- [34] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. arXiv: 13126114.
- [35] Michelucci U. An introduction to autoencoders. arXiv preprint arXiv: 220103898 2022.
- [36] Larsen ABL, Sønderby SK, Larochelle H, et al. Autoencoding beyond pixels using a learned similarity metric. In: International conference on machine learning. PMLR; 2016. p. 1558–66.
- [37] Chen X, Kingma DP, Salimans T, et al. Variational lossy autoencoder. arXiv preprint arXiv:161102731 2016.
- [38] Vahdat A, Kautz J. NVAE: a deep hierarchical variational autoencoder. *Adv Neural Inf Process Syst* 2020;33:19667–79.
- [39] Van Den Oord A, Vinyals O. Neural discrete representation learning. *Adv Neural Inf Process Syst* 2017;30.
- [40] Dupont E. Learning disentangled joint continuous and discrete representations. *Adv Neural Inf Process Syst* 2018;31.
- [41] Razavi A, Van den Oord A, Vinyals O. Generating diverse high-fidelity images with vqvae-2. *Adv Neural Inf Process Syst* 2019;32.
- [42] Higgins I, Matthey L, Pal A, et al. beta-vae: learning basic visual concepts with a constrained variational framework. 2016.
- [43] Kim H, Mnih A. Disentangling by factorising. In: International conference on machine learning. PMLR; 2018. p. 2649–58.
- [44] Chen RT, Li X, Grosse RB, et al. Isolating sources of disentanglement in variational autoencoders. *Adv Neural Inf Process Syst* 2018;31.
- [45] Gomez-Bombarelli R, Wei JN, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent Sci* 2018;4(2):268–76.
- [46] Noh J, Kim J, Stein HS, et al. Inverse design of solid-state materials via a continuous representation. *Matter* 2019;1(5):1370–84.
- [47] Sardeshmukh A, Reddy S, Gautham B, et al. TextureVAE: learning interpretable representations of material microstructures using variational autoencoders. In: AAAI spring symposium. MLPS; 2021.
- [48] Oubari F, De Mathelin A, Décatore R, et al. A binded VAE for inorganic material generation. arXiv preprint arXiv:211209570 2021.
- [49] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Adv Neural Inf Process Syst* 2020;33:6840–51.
- [50] Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis. *Adv. Neural Inform. Process. Sys.* 34 2021;34(Neurips 2021):8780–94.
- [51] Song J, Meng C, Ermon S. Denoising diffusion implicit models. arXiv preprint arXiv:201002502 2020.
- [52] Watson D, Ho J, Norouzi M, et al. Learning to efficiently sample from diffusion probabilistic models. arXiv preprint arXiv:210603802 2021.
- [53] Watson D, Chan W, Ho J, et al. Learning fast samplers for diffusion models by differentiating through sample quality. In: International conference on learning representations; 2022.
- [54] Dockhorn T, Vahdat A, Kreis K. GENIE: higher-order denoising diffusion solvers. arXiv preprint arXiv:221005475 2022.
- [55] Meng C, Gao R, Kingma DP, et al. On distillation of guided diffusion models. arXiv preprint arXiv:221003142 2022.
- [56] Vahdat A, Kreis K, Kautz J. Score-based generative modeling in latent space. *Adv Neural Inf Process Syst* 2021;34:11287–302.
- [57] Austin J, Johnson DD, Ho J, et al. Structured denoising diffusion models in discrete state-spaces. *Adv Neural Inf Process Syst* 2021;34:17981–93.
- [58] Nichol AQ, Dhariwal P. Improved denoising diffusion probabilistic models. In: International conference on machine learning. PMLR; 2021. p. 8162–71.
- [59] Kingma D, Salimans T, Poole B, et al. Variational diffusion models. *Adv Neural Inf Process Syst* 2021;34:21696–707.
- [60] Song Y, Durkan C, Murray I, et al. Maximum likelihood training of score-based diffusion models. *Adv Neural Inf Process Syst* 2021;34:1415–28.
- [61] Bao F, Li C, Zhu J, et al. Analytic-dpm: an analytic estimate of the optimal reverse variance in diffusion probabilistic models. arXiv preprint arXiv: 220106503 2022.
- [62] Lin HJ, Lee K-H, Yun CJ. Microstructure design of multifunctional particulate composite materials using conditional diffusion models. arXiv preprint arXiv:230109051 2023.
- [63] Anand N, Achim T. Protein structure and sequence generation with equivariant denoising probabilistic models. arXiv preprint arXiv: 220515019 2022.
- [64] Schneuing A, Du Y, Harris C, et al. Structure-based drug design with equivariant diffusion models. arXiv preprint arXiv:221013695 2022.
- [65] Shi C, Wang C, Lu J, et al. Protein sequence and structure co-design with equivariant translation. arXiv preprint arXiv:221008761 2022.
- [66] Dinh L, Krueger D, Bengio Y. Nice: non-linear independent components estimation. arXiv preprint arXiv:14108516 2014.
- [67] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp. arXiv:160508803.
- [68] Kingma Durk P, Dhariwal Prafulla. Glow: generative flow with invertible 1x1 convolutions. *Adv Neural Inf Process Syst* 2018;31:10236–45.
- [69] Chen Ricky TQ, Rubanova Yulia, Bettencourt Jesse, Duvenaud David K. Neural ordinary differential equations. *Adv Neural Inf Process Syst* 2018;31: 6572–83.
- [70] Will Grathwohl, Ricky T. Q. Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. FFJORD: free-form continuous dynamics for scalable reversible generative models. arXiv:181001367.
- [71] Ohno Hiroshi. Training data augmentation: an empirical study using generative adversarial net-based approach with normalizing flow models for materials informatics. *Appl Soft Comput* 2020;86:105932.
- [72] Radford Alec, Wu Jeffrey, Child Rewon, Luan David, Amodei Dario, Sutskever Ilya. Language models are unsupervised multitask learners. OpenAI blog 2019;1:9.
- [73] Brown Tom, Mann Benjamin, Ryder Nick, Subbiah Melanie, Kaplan Jared D, Dhariwal Prafulla, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst* 2020;33:1877–901.
- [74] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv:181004805.
- [75] Tong Xie, Yuwei Wan, Wei Huang, Yufei Zhou, Yixuan Liu, and Qingyuan Linghu et al. Large Language models as master key: unlocking the secrets of materials science with GPT. arXiv:230402213.
- [76] Hong Zijian. ChatGPT for Computational Materials Science: A Perspective. *Energy Mater Adv.* 2023;4:0026.
- [77] Chen Xing, Abreu Araujo Flavio, Riou Mathieu, Jacob Torrejon, Ravelosona Dafine, Wang Kang, et al. Forecasting the outcome of spintronic experiments with neural ordinary differential equations. *Nat Commun* 2022;13:1016.
- [78] Li Yajie, Sha Liting, Lv Peili, Qiu Na, Zhao Wei, Chen Bin, et al. Influences of separator thickness and surface coating on lithium dendrite growth: a phase-field study. *Materials* 2022;15:7912.
- [79] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, and Sebastian Borgeaud et al. Emergent abilities of large language models. arXiv: 220607682.
- [80] Liu Yue, Zou Xinxin, Yang Zhengwei, Shi Siqi. Machine learning embedded with materials domain knowledge. *J Chin Ceram Soc* 2022;50:863–76.
- [81] Fuhr Addis S, Sumpter Bobby G. Deep generative models for materials discovery and machine learning-accelerated innovation. *Frontiers in Materials* 2022;9:865270.
- [82] Liu Yue, Yang Zhengwei, Zou Xinxin, Ma Shuchang, Liu Dahui, Avdeev Maxim, et al. Data quantity governance for machine learning in materials science. *Natl Sci Rev* 2023;10. nwad125.
- [83] Shi Siqi, Sun Shiyu, Ma Shuchang, Zou Xinxin, Qian Quan, Liu Yue. Detection method on data accuracy incorporating materials domain knowledge. *J Inorg Mater* 2022;37:1311.
- [84] Weston L, Tshitoyan V, Dagdelen J, Kononova O, Trewartha A, Persson KA, et al. Named entity recognition and normalization applied to large-scale information extraction from the materials science literature. *J. Chem. Inf. Model.* 2019;59:3692–702.
- [85] Tshitoyan Vahe, Dagdelen John, Weston Leigh, Dunn Alexander, Rong Ziqin, Kononova Olga, et al. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* 2019;571:95–8.
- [86] Nie Zhiwei, Zheng Shisheng, Liu Yuanji, Chen Zhefeng, Li Shunning, Lei Kai, et al. Automating materials exploration with a semantic knowledge graph for Li-ion battery cathodes. *Adv Funct Mater* 2022;32:2201437.
- [87] Liu Yue, Ge Xianyuan, Yang Zhengwei, Sun Shiyu, Liu Dahui, Avdeev Maxim, et al. An automatic descriptors recognizer customized for materials science literature. *J Power Sources* 2022;545:231946.

- [88] Liu Yue, Ding Lin, Yang ZhengWei, Ge XianYuan, Liu DaHui, Liu Wei, et al. Domain knowledge discovery from abstracts of scientific literature on Nickel-based single crystal superalloys. *Sci China Technol Sci* 2023. <https://doi.org/10.1007/s11431-022-2283-7>.
- [89] Ishikawa Atsushi. Heterogeneous catalyst design by generative adversarial network and first-principles based microkinetics. *Sci Rep* 2022;12:11657.
- [90] Shen Yujun, Yang Ceyuan, Tang Xiaoou, Zhou Bolei. Interfacegan: interpreting the disentangled face representation learned by GANs. *IEEE Trans Pattern Anal Mach Intell* 2020;44:2004–18.
- [91] Liu Yue, Liu Zitu, Li Shuang, Yu Zhenyao, Guo Yike, Liu Qun, et al. Cloud-VAE: variational autoencoder with concepts embedded. *Pattern Recogn* 2023;140: 109530.
- [92] Laura von Rueden, Sebastian Mayer, Katharina Beckh, Bogdan Georgiev, Sven Giesselbach, and Raoul Heese et al. Informed machine learning—A taxonomy and survey of integrating knowledge into learning systems. *arXiv*: 190312394.
- [93] Zhao Yong, Al-Fahdi Mohammed, Hu Ming, Edirisuriya M, Siriwardane D, Song Yuqi, Nasiri Alireza, et al. High-throughput discovery of novel cubic crystal materials using deep generative neural networks. *Adv Sci* 2021;8: e2100566.
- [94] Ren Pengzhen, Xiao Yun, Chang Xiaojun, Huang Po-yao, Li Zhihui, Chen Xiaojiang, et al. A comprehensive survey of neural architecture search: challenges and solutions. *ACM Comput Surv* 2021;54:1–34.
- [95] Liu Yue, Zhu Kai, Liu Zitu. SSRNAS: search space reduced one-shot NAS by a recursive attention-based predictor with cell tensor-flow diagram. In: 2021 international joint conference on neural networks; 2021. p. 1–8.
- [96] Chen Chun-Teh, Gu Grace X. Generative deep neural networks for inverse materials design using backpropagation and active learning. *Adv Sci* 2020;7: 1902607.
- [97] Jonathan Frankle, and Michael Carbin. The lottery ticket hypothesis: finding sparse, trainable neural networks. *arXiv*:180303635.
- [98] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale GAN training for high fidelity natural image synthesis. *arXiv*:180911096.
- [99] Ganguli Deep, Hernandez Danny, Lovitt Liane, Askell Amanda, Bai Yuntao, Chen Anna, et al. Predictability and surprise in large generative models. In: 2022 ACM conference on fairness, accountability, and transparency; 2022. p. 1747–64.
- [100] Hu Hailong, Pang Jun. Stealing machine learning models: attacks and countermeasures for generative adversarial networks. In: Annual computer security applications conference; 2021. p. 1–16.



Yue Liu obtained her B.S. and M.S. in computer science from Jiangxi Normal University in 1997 and 2000. She finished her Ph.D. in control theory and control engineering from Shanghai University (SHU) in 2005. She has been working with the School of Computer Engineering and Science of SHU since July 2000. During that time, she has been a curriculum R&D manager at the Sybase-SHU IT Institute of Sybase Inc. from July 2003 to July 2004 and a visiting scholar at the University of Melbourne from Sep. 2012 to Sep. 2013. At present, she is professor of SHU. Her current research interests focus on research of data mining, machine learning, and AI for materials science.



Siqi Shi obtained his B.S. and M.S. from Jiangxi Normal University in 1998 and in 2001, respectively. He finished his Ph.D. from Institute of Physics, Chinese Academy of Sciences in 2004. During this period, focused on the electrolyte, electrode materials and relevant interfaces for lithium-ion batteries, he carried out the first-principles calculation and design on the ionic transport physics, cooperative electron/ion transport control problem earliest in China. After that, he joined the National Institute of Advanced Industrial Science and Technology of Japan and Brown University of USA as a senior research associate until joining Shanghai University as a professor in early 2013. His current research interests focus on the fundamentals and multiscale calculation of electrochemical energy storage materials and materials design and performance optimization using machine learning.