Matter

CellPress
OPEN ACCESS

Perspective

Has generative artificial intelligence solved inverse materials design?

Hyunsoo Park,[1] Zhenzhu Li,[1,2] and Aron Walsh[1,3,*]

SUMMARY

The directed design and discovery of compounds with pre-determined properties is a long-standing challenge in materials research. We provide a perspective on progress toward achieving this goal using generative models for chemical compositions and crystal structures based on a set of powerful statistical techniques drawn from the artificial intelligence community. We introduce the central concepts underpinning generative models of crystalline materials. Coverage is provided of early implementations for inorganic crystals based on generative adversarial networks and variational autoencoders through to ongoing progress involving autoregressive and diffusion models. The influence of the choice of chemical representation and the generative architecture is discussed, along with metrics for quantifying the quality of the hypothetical compounds produced. While further developments are required to enable realistic predictions drawn from richer structure and property datasets, generative artificial intelligence is already proving to be complementary to traditional materials design strategies.

INTRODUCTION

The most common workflows in materials modeling are based on the structure → property paradigm.[1] Such a mapping can be achieved with an appropriate property calculator. This may involve a numerical solution to a quantum mechanical expression (e.g., Kohn-Sham equations),[2] the use of an analytical interatomic force field (e.g., Buckingham potential),[3] or the training of a surrogate machine learning model (e.g., crystal graph convolutional neural network [CGCNN]).[4]

In a traditional high-throughput screening approach, a pool of candidate materials (σ_{pool}) is filtered by a series of hand-built criteria. The aim is to identify the optimal target configurations (σ_{target}), whose properties exceed a certain threshold P_{target}. A high-throughput screening process is illustrated in Figure 1. While this approach has proved successful for diverse application areas, from thermoelectrics to batteries,[5] it is often limited to a particular set of crystal structures and/or chemical compositions. The identified targets are a subset of the input pool of candidate compounds, i.e., $\sigma_{target} \subseteq \sigma_{pool}$.

The inverse approach involves mapping from property → structure.[6] Rather than starting with a known set of material structures and predicting their properties, the process begins by articulating a desired set of properties (constraints) and seeks to identify the corresponding molecular or extended crystal structures that satisfy them. This change in direction widens the chemical search space but also increases the complexity of the problem, prompting the development and adoption of new computational techniques and methodologies. Recent advances in

PROGRESS AND POTENTIAL

Computer simulations serve as a powerful tool for investigating the relationship between the composition, structure, and properties of materials. Early approaches simplified the many-body interactions inherent in real systems into tractable functional forms, solvable either manually or with calculators. High-performance computing ushered in an era where numerical methods could tackle increasingly complex equations, with density functional theory becoming a dominant first-principles technique for materials modeling. Now, artificial intelligence (AI) provides new opportunities for statistical descriptions of many types of materials drawn from the large volume of data that have been produced. Building on the use of machine learning surrogate models that enable rapid property predictions for materials, generative AI can enable researchers to actively propose entirely new structures or compounds that may exhibit desired characteristics. In this perspective, we introduce some of the central concepts underpinning generative AI, with a focus on its application to inorganic crystals. The aim is for materials scientists to grasp the terminology, how these techniques work, what models are available, and opportunities for using generative AI in materials research.

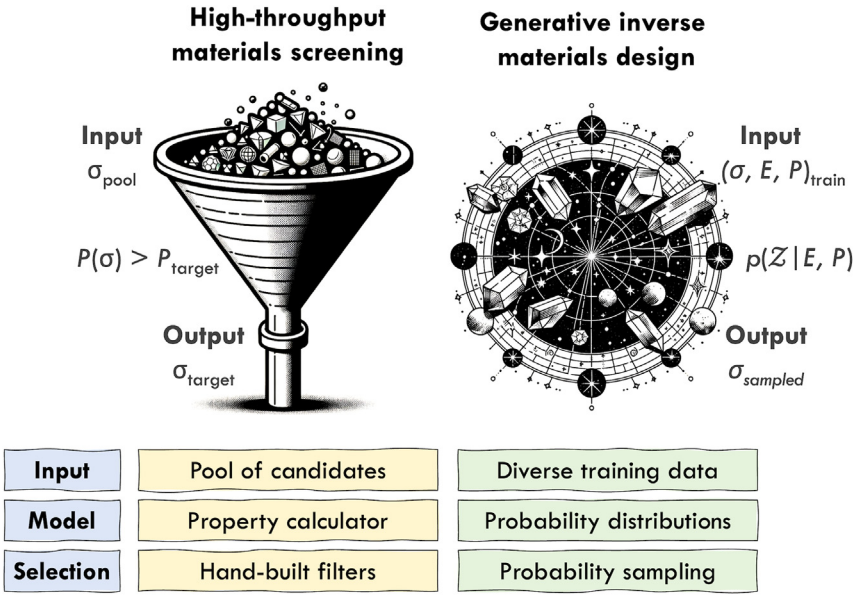

Figure 1. Comparison of traditional high-throughput screening and generative AI approaches
The inset images were generated using a text-to-image AI model.[10] The configuration, energy, and properties of materials are indicated by σ, *E*, and *P*, respectively.

materials datasets and machine learning have been instrumental in making the inverse design of materials feasible. Generative techniques, which leverage probabilistic models to create novel data, have become popular in the artificial intelligence (AI) community for solving inverse problems across various domains such as natural language processing, audio, image synthesis,[7] and recent advances in video generation.[8]

Due to their potential for targeted molecular and materials design, generative techniques have been described as "among the greatest opportunities available in modern chemical research."[9] Actively explored generative models and outstanding challenges for molecular generative research are reviewed in Anstine and Isayev.[9] As rapid progress is being made in the development and usage of generative techniques for crystalline inorganic materials, the purpose of this perspective is to introduce the key concepts, advantages, and limitations of these applications. In particular, we delve into some early implementations, summarize emerging trends in recent implementations, and highlight new directions for exploration.

HOW DO GENERATIVE MODELS WORK?

The first generative models involved a combination of three processes to learn and synthesize materials data: encoding, decoding, and generation.

(1) Encoding is the process of transforming the configuration (σ) of a material into a different representation. The input (training) data are in the form of a numerical tensor that may encompass the chemical composition, crystal structure, and relevant physical properties of each system. Encoding compresses the input data into lower dimensions, referred to as a latent or feature space (Z).

$$Z = f(\sigma) \qquad \text{(Equation 1)}$$

[1]Thomas Young Centre and Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, UK

[2]I-X Centre for AI in Science, Imperial College London, White City Campus, 84 Wood Lane, London W12 0BZ, UK

[3]Department of Physics, Ewha Womans University, Seoul 03760, Korea

*Correspondence: a.walsh@imperial.ac.uk

https://doi.org/10.1016/j.matt.2024.05.017

For crystalline materials, the standard configuration σ is defined by the atom types or chemical composition ($\mathcal{A}$), lattice vectors of the unit cell ($\mathcal{L}$), and the atomic coordinates within a 3-dimensional (3D) Euclidean space ($\mathcal{X}$). The function f for transforming these representations into the latent space has no fixed form and is learned through a deep learning architecture such as a geometric graph neural network (GNN).[11]

(2) Decoding is the inverse process of encoding. It involves transforming the encoded or latent representation back into the original data space using a learned function (f'). The decoder network in generative models is responsible for this reconstruction. When decoded successfully, the composition, structure, and properties of the generated material can be obtained.

$$\sigma' = f'(Z) \quad \text{(Equation 2)}$$

The unavoidable losses that occur by encoding and decoding result in noise (reconstruction error) that can cause issues such as spurious changes in symmetry or stoichiometry. These losses sometimes lead to the generation of invalid structures (e.g., missing atoms), worsened by errors in sampling that undermine the fidelity and reliability of the sampled compounds.

(3) Generation is the creative process where novel compounds are produced. This goal is achieved by sampling from a probability distribution in the learned latent space. For example, the likelihood of finding atoms at different positions within a unit cell can be described as a probability distribution. The novel instances produced ($\sigma_{sampled}$)) should share characteristics with the training data, which may have been a set of structures with their corresponding labeled energies and properties, i.e., $(\sigma, E, P)_{train}$ in Figure 1. The effective sampling of new configurations from the joint probability distribution is crucial for creating diverse and realistic samples, contributing to the ability of a model to produce chemically sensible and original outputs.

$$\sigma_{sampled} \sim p(Z) \quad \text{(Equation 3)}$$

Recent advancements in generative models now incorporate a broader array of techniques, including the direct sampling of materials from noise and responding to user text prompts. Four generative architectures that have been applied to inorganic crystals, illustrated in Figure 2, are (1) variational autoencoders (VAEs) that combine an encoder to a continuous learned latent space with a decoder to generate samples by drawing from the probability distribution; (2) generative adversarial networks (GANs) based on a generator model for synthetic data (e.g., sampled structures), which tries to fool a discriminator model that distinguishes between real and synthetic data; (3) diffusion models that generate sampled materials through a series of iterative stochastic transformations applied to an initial (noisy) data distribution; and (4) autoregressive models, including transformer-based large language models (LLMs) that sample the conditional probability distribution of a sequence, leveraging the captured data dependencies to sequentially generate hypothetical compounds.

HOW IS INVERSE MATERIALS DESIGN ADDRESSED?

Analogs to the three processes outlined above can be found in traditional materials modeling approaches. For example, crystal structure prediction aims to minimize the energy of a chemical system with respect to the global configuration of atoms. Genetic algorithms can be used for this purpose, where the structure of a chemical

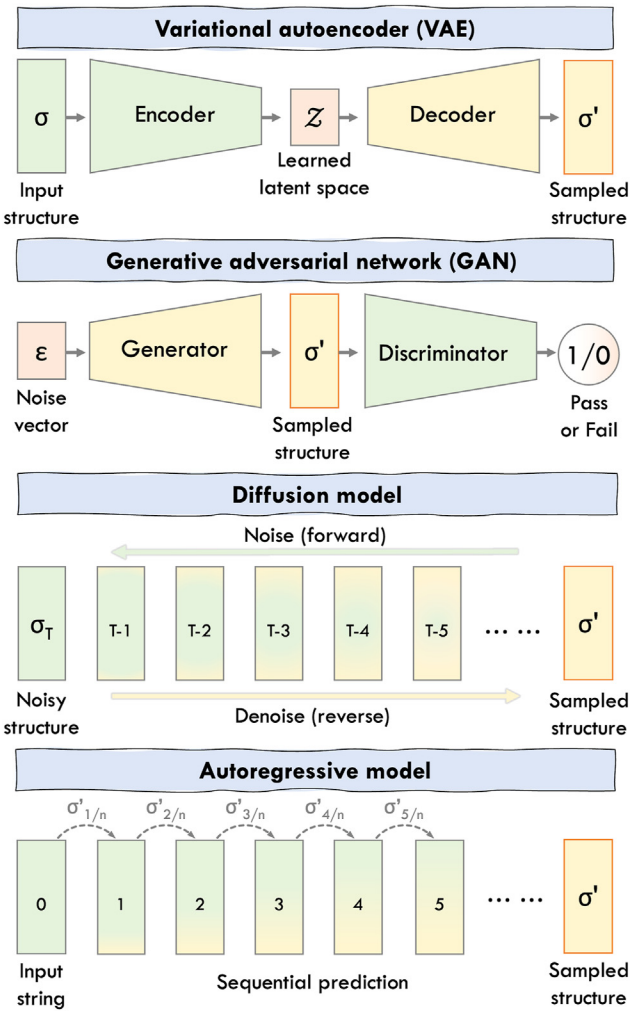

Figure 2. Illustration of generative model architectures

Four models that can be tailored for materials design and discovery ranging from the sampling of latent space in a VAE to text sequence modification in an autoregressor. Each model involves a distinct deep learning approach that can generate new materials. An ideal model will combine fast sampling of structures and properties with high quality and diversity.

system is encoded as a string of 1s and 0s.[12] These are mutated to generate (explore) new configurations and decoded back into a standard structure representation. A common use for this type of algorithm is to efficiently identify the global minimum configuration:

$$\sigma_{target} = \min_{\sigma}[E(\sigma)] \equiv \min_{Z}[E(Z)] \quad \text{(Equation 4)}$$

with equivalent solutions being found for optimization in real space or latent space. Similarly, property optimization can be performed by iteratively adjusting the configuration of the system toward a target property or weighted set of properties (P_{target}) across the multi-objective Pareto front, e.g.,

$$\sigma_{target} = \min_{Z}\left|P(Z) - P_{target}\right| \quad \text{(Equation 5)}$$

which has been used for finding cases as diverse as direct band-gap polymorphs of Si[13] to perovskite structured crystals suitable for water splitting.[14]

Standard high-throughput materials screening and design techniques can require thousands of individual calculations to identify target compounds within limited chemical spaces. In contrast, a modern generative approach may directly produce a solution that is sampled from the learned latent space and conditioned to satisfy particular thermodynamic and property objectives:

$$\sigma_{sampled} \sim p\left(Z|E_{\text{target}}, P_{\text{target}}\right) \quad \text{(Equation 6)}$$

This conditional probability distribution represents the probability of a compound being drawn from Z given specific characteristics such as target energy and electronic properties. Of course, there is no free lunch in terms of the required computational cost. The hidden overhead for generative models is in the training and testing, which requires a diverse dataset to define the initial distribution of compounds and a significant computing resource for parameterizing complex models. The first obstacle is being alleviated by an increasing volume of high-quality materials data available in public datasets,[15,16] whereas the second is being addressed with national public CPU supercomputers being augmented with GPU capabilities, driven by the growing importance and utility of deep learning approaches.

EVOLUTION OF GENERATIVE MODELS FOR MATERIALS

Dozens of generative models have been reported for materials, which involve some aspects of composition, structure and/or property sampling. A selection of published models for inorganic crystals is listed in Table 1. These involve different combinations of structure representation and generation architecture. Open-source code repositories are available for the majority of these models, which has supported rapid progress and collaboration in the field. However, we note that the documentation may be limited, and several codes are not actively developed and may not run with the latest compilers or deep learning libraries.

Structure representation and encoding

One aspect that can distinguish between models is the choice of crystal structure representation. CrystalGAN, reported in 2018,[17] was based on a standard crystallographic representation of a 3×3 matrix describing the x, y, and z components of the a, b, and c lattice vectors of the unit cell ($\mathcal{L}$), in addition to fractional coordinates ($\mathcal{X}$) for atom types ($\mathcal{A}$). While convenient, this numerical representation lacks invariance with respect to relevant symmetry operations, i.e., the feature values representing a given structure are changed by unit cell translations, rotations, and atomic permutations.[34] This is also true for the iMatGen model, published in 2019,[19] which represents the unit cell of a crystal using voxels (a 3D grid with a spacing of 0.23 Å). While there are data augmentation tricks to alleviate the lack of invariance (e.g., training data containing shuffled atomic positions, as well as translated and rotated versions of unit cells), this approach does not solve the problem of representing crystal materials effectively.

Alternatively, it is more rigorous and efficient to encode the structures in a way that respects the relevant functional symmetry from the outset, such as a crystal graph.[34] GNNs stand out in this regard, offering a solution by abstracting information from the standard crystallographic representation while ensuring invariance to key transformations such as permutations, translations, and rotations. For crystalline materials, it is also important to consider periodic boundary conditions that describe the repeating nature of the unit cells. The multi-graph representation, introduced by CGCNN[4] provided an important step by representing atoms as nodes and the connections (or distances) between neighboring atoms as edges, considering finite

Table 1. Selection of generative models developed for crystalline inorganic materials

Model	Structure representation	Generation architecture	License	Reference
CrystalGAN[a]	coordinates	generative adversarial network (GAN)	GPL	Nouira et al.[17]
Hoffmann et al.[b]	voxels	variational autoencoder (VAE)	MIT	Hoffmann et al.[18]
iMatGen[c]	voxels	VAE	–	Noh et al.[19]
CCCGAN[d]	coordinates	GAN	–	Kim et al.[20]
ICSG3D[e]	voxels	VAE	MIT	Court et al.[21]
CCDCGAN[f]	voxels	GAN	–	Long et al.[22]
CubicGAN[g]	coordinates	GAN	MIT	Zhao et al.[23]
FTCP[h]	coordinates and structure factors	VAE	Apache	Ren et al.[24]
CDVAE[i]	coordinates	VAE + diffusion	MIT	Xie et al.[25]
CrystaLLM[j]	crystallographic information file	autoregressive transformers (ATs)	MIT	Antuneset al.[26]
xyzTransformer[k]	crystallographic information file	ATs	–	Flam-Shepherdet al.[27]
CrysTens[l]	coordinates	GAN, diffusion	MIT	Alverson et al.[28]
DiffCSP[m]	coordinates	diffusion	MIT	Jiao et al.[29]
MatterGen[n]	coordinates	diffusion	–	Zeni et al.[30]
WyCryst[o]	Wyckoff site matrix	VAE	–	Zhu et al.[31]
UniMat[p]	coordinate embedded periodic table	diffusion	–	Yang et al.[32]
Crystal-llm[q]	coordinates	ATs	CC-BY-NC	Gruver et al.[33]

The chronological list is not comprehensive but covers relevant early models and some major recent developments.
[a]https://github.com/asmanouira/CrystalGAN.
[b]https://github.com/hoffmannjordan/Encoding-Decoding-3D-Crystals.
[c]https://github.com/kaist-amsg/imatgen.
[d]https://github.com/kaist-amsg/Composition-Conditioned-Crystal-GAN.
[e]https://github.com/by256/icsg3d.
[f]https://github.com/TengLong1993/CCDCGAN-for-single-system.
[g]https://github.com/MilesZhao/CubicGAN.
[h]https://github.com/PV-Lab/FTCP.
[i]https://github.com/txie-93/cdvae.
[j]https://github.com/lantunes/CrystaLLM.
[k]https://github.com/danielflamshep/xyztransformer.
[l]https://github.com/michaeldalverson/CrysTens.
[m]https://github.com/jiaor17/DiffCSP.
[n]None yet.
[o]https://github.com/RaymondZhurm/WyCryst.
[p]None yet.
[q]https://github.com/facebookresearch/crystal-text-llm.

cutoffs beyond the contents of an isolated unit cell. Crystal Diffusion Variational Autoencoder (CDVAE), reported in 2022,[25] then combined invariant encoding of $\mathcal{A}$, $\mathcal{L}$, and $\mathcal{X}$ with an SE(3)-equivariant GNN. This choice ensures that the model leverages physical symmetries, leading to a richer representation and transformation of structure, while also requiring fewer data.

Foundational generation architecture

Drawing from developments in the generative AI field, the underlying model architectures have also evolved in the materials domain. The first models were based on GANs and VAEs.

GANs operate through the dynamic interplay between a generator and a discriminator that is trained to distinguish between real and fake material data. The adversarial process fosters a continuous enhancement of the generator's capability as it learns to approximate the distribution of the materials in the training set. In 2018, CrystalGAN leveraged GANs to design ternary A-B-H compositions, generating 3D coordinates ($\mathcal{X}$) and lattice vectors ($\mathcal{L}$). This model applied cross-domain learning to sample ternary materials from binary compounds. Composition Conditioned Crystal GAN

(CCCGAN), published in 2020,[20] employed composition embedding vectors as inputs for generating composition-conditioned crystals with GANs. Constrained Crystals Deep Convolutional GAN (CCDCGAN), published in 2021,[22] created a 3D voxel representation of materials built on a deep convolutional GAN. However, this approach faced challenges such as the need for post-processing requirements and memory-intensive input data representation. Additionally, CubicGAN, published in 2021,[23] integrated space group information into the training of GANs for the generation of novel prototype materials within the cubic crystal system.

GANs implicitly learn the data distribution using a generator and a discriminator, while VAEs explicitly model the data distribution by mapping it onto known distributions with an encoder and a decoder. VAEs combine an encoder that maps the structure representation into latent space and a decoder that reconstructs them. The training of VAEs involves optimizing a loss function composed of two terms: reconstruction loss and regularization loss. Reconstruction measures the fidelity with which the decoder can reconstruct the input data from their latent representation. Regularization encourages the latent space to adhere to a specified distribution, typically Gaussian. This regularization should facilitate the generation of new data points (compounds) that are variations of the input data by promoting a well-structured latent space that supports effective sampling and interpolation. iMatGen leveraged a VAE architecture to generate voxel representations of crystal materials.[19] The model first compresses the 3D image representation of crystal structures. The decoder is trained not only on the standard VAE loss but also to classify their stability based on the labeled formation energy. These objectives, in principle, allow the model to navigate materials space more effectively, guiding it toward thermodynamically stable compounds.

Fourier Transformed Crystal Properties (FTCP), published in 2022,[24] combined a VAE architecture with an invertible crystallographic representation. FTCP integrates both real-space and reciprocal-space features drawn from crystallography. However, the representation scheme does not satisfy crystal invariances, and practically, the model suffers from heavy reconstruction losses, leading to frequent invalid structures where the geometry is wrong or subsequent structural relaxation fails.[24,25] WyCryst, reported in 2023,[31] introduced a Wyckoff site matrix that includes the occupancy of each of the Wyckoff positions in the unit cell. This invertible representation is encoded in a VAE based on a convolutional neural network and demonstrates lower reconstruction losses and higher structure validity than FTCP. The discretized description of atomic positions may be most appropriate for high-symmetry prototype crystal structures that avoid high internal degrees of freedom.

Both VAEs and GANs have been widely used for materials generation but have known limitations. The adversarial training process inherent in GANs is often unstable, which has led to extensions of the "vanilla flavor" with modified loss functions such as the Wasserstein (Earth-Mover's) distance or a least squares loss. One study on the generation of chemical compositions within the elpasolite (double perovskite) structure type compared both architectures and showed that the VAE model showed slightly higher precision and was easier to train than a Wasserstein GAN.[35] A common problem with early generative models is pathologies that can ignore data variations (VAEs) or produce indistinguishable outputs (GANs).[36] Such issues that limit the diversity of the generated data have motivated the development of new architectures, as well as alternative generative modeling strategies.[37]

Diffusion era

In materials science, first-order phase transitions (e.g., solid/liquid) are discontinuous and disruptive, whereas second-order transitions (e.g., between certain magnetic configurations) are continuous and reversible. There is an analogy here: the underlying data transformations in GANs and VAEs can be seen as a first order process, while diffusion models reformulate data generation into a smoother, second-order process.

Diffusion models are based on a step-by-step data transformation that aids model training.[38] They operate through a forward process that incrementally adds noise to an original data point (compound) across multiple steps, converting it into pure noise, and a reverse process that reconstructs new data samples (compounds) from this noisy state. This strategy has been discussed in the context of symmetry breaking in statistical physics.[39]

CDVAE, published in 2022,[25] integrates a diffusion model within VAEs, specifically replacing the decoder of VAE with a noise conditional score network (NCSN).[40] This model is trained on stable materials, enabling the sampling of new material structures by gradually removing noise through Langevin dynamics, starting from noisy data and guiding the reconstruction toward stable configurations.

DiffCSP, reported in 2023,[29] showcased the capability of diffusion models for crystal generation. By jointly applying a diffusion process with lattice vectors ($\mathcal{L}$) and fractional coordinates ($\mathcal{X}$), the model can better capture the geometry of crystals, ensuring that the generated structures adhere to physical and chemical expectations. This approach is distinct from CDVAE, which first predicts $\mathcal{L}$ and then updates $\mathcal{X}$.

MatterGen, reported in 2023,[30] demonstrated that diffusion models can flexibly adapt to generate materials with desired requirements such as composition, symmetry, and mechanical properties using a classifier-free guidance algorithm.[41] To pre-train MatterGen, a large dataset was constructed of approximately 1 million unique bulk crystal structures sourced from multiple databases. This dataset empowers the learned distributions and the generation of stable and varied materials. Notably, the approach adopted by MatterGen allows for the generation of materials that meet multiple property constraints simultaneously by employing further fine-tuning.

Another diffusion model, UniMat, was reported in 2023.[32] It used a novel representation scheme that stores traditional crystal coordinates at the corresponding element entry in the periodic table, i.e., including an explicit index of chemical groups and periods. This choice leads to a sparse tensor representation where a null value of -1 indicates the absence of an atom type in a given material. Data augmentation is used to account for the lack of crystal invariance in this representation. Conditional diffusion is again performed using the classifier-free guidance introduced by Ho and Salimans.[41]

The effectiveness of diffusion models for materials generation has been demonstrated in several comparative studies. For example, the learned distribution of crystal structure parameters was much better for a diffusion model based on CrysTens compared to both a regular GAN and a Wasserstein GAN.[28] Similarly, in the ZeoDiff generator for zeolitic materials, diffusion greatly outperforms a GAN in terms of the structure validity of the sampled compounds.[42]

Language of materials

Autoregressive models generate data sequentially, conditioning each prediction on previous outputs. Autoregressive LLMs have emerged as a popular and powerful branch of AI. Their natural language processing capabilities can be leveraged for materials science in several ways, such as data extraction[43] and problem solving.[44] They also present an alternative avenue for materials generation. Compared to recurrent neural networks, autoregressive models based on the transformer architecture excel in capturing long-range dependencies.[45] Despite the complex nature of the models and the substantial training data they require, there have been successful attempts to generate materials in text format.

One application is the generation of text files. CrystaLLM and xyzTransformer, both reported in 2023,[26,27] employ GPT-type models, predominantly the decoder part of transformer architectures, trained to explicitly generate crystallographic information files (CIFs), among other file formats. To overcome the inherent lack of invariance in the text files when encoding crystal structure information, data augmentation techniques involving randomly rotated structures were utilized.

Instead of training a model from scratch, Crystal-llm, reported in 2024,[33] fine-tuned LLAMA-2.[46] This approach not only streamlined the process but also showcased high performance in sampling inorganic compounds, surpassing the first wave of diffusion-based models such as CDVAE in targeting low-energy configurations. Furthermore, the potential of leveraging text prompts for conditionally generating crystals was highlighted, underscoring the adaptability of autoregressive models.

TURING TEST FOR MATERIALS GENERATION

Materials scientists are trained to understand the relationship between composition, structure, and properties. This includes knowing when a certain chemical composition will be synthetically accessible, a crystal structure is implausible, or a physical property may be unrealistic. It could be argued that there is an inherent bias in how materials scientists think, informed by empirical rules and heuristics. A statistical approach including non-linear correlations that are difficult to distil into simple teachable principles may produce highly performant unexpected solutions, especially when trained on diverse datasets.

The true utility of generative materials models will be realized when they can perform as well as or better than human experts. In the spirit of the Turing test, the goal could be achieved when AI-generated materials are indistinguishable from human-designed ones or are clearly made by machines because of their exceptional structures and/or properties. This is a high bar to achieve. Even traditional knowledge-led computational materials design is sometimes criticized for predictions that are obvious (trivial variants of known systems) or erroneous (compounds that are unstable or fail to exhibit the predicted structure/properties), e.g., as discussed recently in the context of autonomous materials synthesis[47] and computational materials discovery.[48] Of course, full autonomy is not essential, and generative models may still play a valuable role in providing novel suggestions while keeping materials scientists and complementary computational search strategies "in the loop."

To illustrate a simple use case for a generative model, we consider the inorganic compound $Sb_5S_4Cl_2$. It is a chemically plausible hypothetical composition for which crystal structures have previously been proposed based on data mining (substitution into known structures) and an evolutionary global search using high-performance

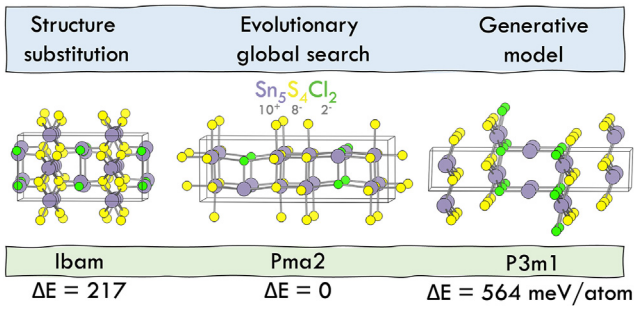

Figure 3. Candidate crystal structures for the inorganic compound $Sn_5S_4Cl_2$

Three locally stable polymorphs were found by substitution into known prototypes,[49] global optimization using an evolutionary algorithm,[49] and a generative LLM for crystals, CrystaLLM.[26] The total energies, calculated using the MACE-MP universal force field,[50] are compared.

computing.[49] Those two crystal structures are compared to one generated by CrystaLLM in Figure 3. The generated structure is chemically reasonable at first glance but, among the three distinct Sn(II) sites, is a rather improbable linear Cl–Sn–Cl environment. As a result, the calculated total energy of this structure is significantly higher than for the other candidates. This comparison is based on a single stochastic sample from the generative model, which has not been conditioned to target low-energy structures.

Beyond structure generation, models can also be trained to target compounds with specific chemistry, symmetry, or properties. Two of the inorganic crystals sampled by MatterGen include K_3AlCl_6, which contains elements in their standard oxidation states, and $NaNiIO_6$, which features a more exotic Ni^{4+} cation and the orthoperiodate (IO_6^{5-}) polyanion. Crystals for both compositions have been reported experimentally.[51,52] Interestingly, by targeting compounds with a high bulk modulus, Re_3B_2C is predicted as a target layered material with no previous literature reports. One restriction is that current models are constrained to ordered crystals with small unit cells owing to the nature of the training data and limitations in the model complexity.

Practically, the field could benefit from a consistent set of metrics for assessing generative performance. For vision-based models, a range of metrics has been used to assess the diversity and quality of generated images, such as the inception score (IS) and the Fréchet inception distance (FID), contributing significantly to the field's advancement.[53] In the context of crystal structures, metrics such as the validity of generated compositions (ensuring atomic arrangements free from overlaps) and the similarity between generated materials and known materials are employed.[25] Additionally, the novelty (or uniqueness) can be quantitatively assessed using an approach such as the "StructureMatcher" function in pymatgen,[54] which compares atomic arrangement based on the reduced chemical formula and space group.[30] Such metrics are being combined in matbench-genmetrics.[55]

While the total energy from first-principles calculations is a useful quantity, the internal energy of a material is a poor indicator of thermodynamic stability or practical synthesizability in the lab, which involves free energies and chemical kinetics and depends on the choice of processing route.[56,57] As interest grows in metastable crystals (kinetically stabilized local solutions), the internal energy will not be sufficient to distinguish poor from promising candidate compounds.

It will be important to develop and incorporate more comprehensive metrics for generated materials that encapsulate not only the structural fidelity, novelty, and diversity but also synthetic routes to realize them. As these methods mature and become more accessible with increasing computer power, blind tests could be developed mirroring those held in the crystal structure prediction community.[58]

OUTLOOK

Generative models offer a route to navigate high-dimensional, non-linear, material spaces that are incomprehensible to human scientists. Progress in the development of generative models of materials has been rapid over the past 6 years and is accelerating. At first glance, the underlying statistical techniques appear to depart from traditional materials modeling techniques. However, digging deeper shows similarities and analogies in the approaches. The choice of how chemistry is represented to construct effective probability distributions in these deep learning models is critical to their success.

One challenge for generative AI is interpretability and the extraction of physical principles. These are large and complex models that can involve millions of parameters and can be used as a black-box approach to make predictions. However, the underlying learned probability distributions contain fundamental composition-structure-property relationships that enable the material design. These can be explored and visualized in different ways to help scientists extend existing physical principles and perhaps to develop new ones.

While predictions of bulk crystal structures and properties have been the focus of the first wave of generative models for materials, they are not limited to this purpose. The sampling of probability distributions may be appropriate for many other combinatorial tasks that are difficult to solve using standard approaches. For example, recent studies reported using generative diffusion models to sample reconstructions for oxide formation on (111) surfaces of Ag[59] and to design porous zeolitic materials.[42] Another study trained a GAN to sample multi-component alloy compositions, directing the experimental synthesis of $Al_5Co_8Cu_{35}Fe_{19}Ni_{23}V_{11}$[60] which would be intractable for standard high-throughput approaches. A similar strategy was reported to generate compositionally complex bulk metallic glasses[61] and the processing parameters to control thin-film microstructure.[62]

There is no barrier to further extending these techniques to other problems where multiple configurations must be accessed. Examples include how the structure and composition of an electrode evolves during the charging cycle of a battery or how the microstructure of a solid catalyst changes during a multi-step reaction cycle. Generative approaches can also be combined with optimization techniques such as Bayesian optimization or reinforcement learning to develop powerful workflows, as demonstrated for small-molecule design.[63] We may be at the proof-of-concept stage where sensible solutions are an achievement rather than predictions that are truly novel or exciting to a materials scientist. There is hope, however, for future generations.

ACKNOWLEDGMENTS

We thank Dylan M. Anstine for critical feedback, as well as Tian Xie and Ekin Dogus Cubuk for stimulating discussions. This work was supported by EPSRC project EP/X037754/1 and the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Sciences program.

AUTHOR CONTRIBUTIONS

H.P. and A.W. conceptualized this project. H.P., Z.L., and A.W. wrote, reviewed, and edited the manuscript. A.W. guided the content of this article. All authors discussed and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used DALL·E 3 to create the inset images in Figure 1. The authors reviewed and edited the figure as needed and take full responsibility for the content of the publication.

REFERENCES

1. Dirac, P.A.M. (1929). Quantum Mechanics of Many-Electron Systems. Proc. Roy. Soc. Lond.: Math. Phys. Eng. Sci. *123*, 714–733.
2. Kohn, W., and Sham, L.J. (1965). Self-Consistent Equations Including Exchange and Correlation Effects. Phys. Rev. *140*, A1133–A1138. https://doi.org/10.1103/PhysRev.140.A1133.
3. Lewis, G.V., and Catlow, C.R.A. (1985). Potential models for ionic oxides. J. Phys. C Solid State Phys. *18*, 1149–1161. https://doi.org/10.1088/0022-3719/18/6/010.
4. Xie, T., and Grossman, J.C. (2018). Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Phys. Rev. Lett. *120*, 145301. https://doi.org/10.1103/PhysRevLett.120.145301.
5. Curtarolo, S., Hart, G.L.W., Nardelli, M.B., Mingo, N., Sanvito, S., and Levy, O. (2013). The high-throughput highway to computational materials design. Nat. Mater. *12*, 191–201. https://doi.org/10.1038/nmat3568.
6. Franceschetti, A., and Zunger, A. (1999). The inverse band-structure problem of finding an atomic configuration with given electronic properties. Nature *402*, 60–63. https://doi.org/10.1038/46995.
7. Harshvardhan, G., Gourisaria, M.K., Pandey, M., and Rautaray, S.S. (2020). A comprehensive survey and analysis of generative models in machine learning. Computer Science Review *38*, 100285.
8. Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., Yuan, Z., Huang, Y., Sun, H., Gao, J., et al. (2024). Sora: A review on background, technology, limitations, and opportunities of large vision models. Preprint at arXiv. https://doi.org/10.48550/arXiv.2402.17177.
9. Anstine, D.M., and Isayev, O. (2023). Generative Models as an Emerging Paradigm in the Chemical Sciences. J. Am. Chem. Soc. *145*, 8736–8750. https://doi.org/10.1021/jacs.2c13467.
10. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., and Chen, M. (2022). Hierarchical Text-Conditional Image Generation with CLIP Latents. Preprint at arXiv. [cs]. https://doi.org/10.48550/arXiv.2204.06125.
11. Duval, A., Mathis, S.V., Joshi, C.K., Schmidt, V., Miret, S., Malliaros, F.D., Cohen, T., Liò, P., Bengio, Y., and Bronstein, M. (2023). A hitchhiker's guide to geometric gnns for 3d atomic systems. Preprint at arXiv. https://doi.org/10.48550/arXiv.2312.07511.
12. Woodley, S.M., and Catlow, R. (2008). Crystal structure prediction from first principles. Nat. Mater. *7*, 937–946. https://doi.org/10.1038/nmat2321.
13. Lee, I.-H., Oh, Y.J., Kim, S., Lee, J., and Chang, K.J. (2016). *Ab initio* materials design using conformational space annealing and its application to searching for direct band gap silicon crystals. Comput. Phys. Commun. *203*, 110–121. https://doi.org/10.1016/j.cpc.2016.02.011.
14. Jain, A., Castelli, I.E., Hautier, G., Bailey, D.H., and Jacobsen, K.W. (2013). Performance of genetic algorithms in search for water splitting perovskites. J. Mater. Sci. *48*, 6519–6534. https://doi.org/10.1007/s10853-013-7448-9.
15. Walsh, A. (2024). Open computational materials science. Nat. Mater. *23*, 16–17. https://doi.org/10.1038/s41563-023-01699-7.
16. Ghiringhelli, L.M., Carbogno, C., Levchenko, S., Mohamed, F., Huhs, G., Lüders, M., Oliveira, M., and Scheffler, M. (2017). Towards efficient data exchange and sharing for big-data driven materials science: metadata and data formats. npj Comput. Mater. *3*, 46. https://doi.org/10.1038/s41524-017-0048-5.
17. Nouira, A., Sokolovska, N., and Crivello, J.-C. (2018). CrystalGAN: Learning to Discover Crystallographic Structures with Generative Adversarial Networks. Preprint at arXiv, 1810. https://doi.org/10.48550/arXiv.1810.11203.
18. Hoffmann, J., Maestrati, L., Sawada, Y., Tang, J., Sellier, J.M., and Bengio, Y. (2019). Data-Driven Approach to Encoding and Decoding 3-D Crystal Structures. Preprint at arXiv. https://doi.org/10.48550/arXiv.1909.00949.
19. Noh, J., Kim, J., Stein, H.S., Sanchez-Lengeling, B., Gregoire, J.M., Aspuru-Guzik, A., and Jung, Y. (2019). Inverse Design of Solid-State Materials via a Continuous Representation. Matter *1*, 1370–1384. https://doi.org/10.1016/j.matt.2019.08.017.
20. Kim, S., Noh, J., Gu, G.H., Aspuru-Guzik, A., and Jung, Y. (2020). Generative Adversarial Networks for Crystal Structure Prediction. ACS Cent. Sci. *6*, 1412–1420. https://doi.org/10.1021/acscentsci.0c00426.
21. Court, C.J., Yildirim, B., Jain, A., and Cole, J.M. (2020). 3-D Inorganic Crystal Structure Generation and Property Prediction via Representation Learning. J. Chem. Inf. Model. *60*, 4518–4535. https://doi.org/10.1021/acs.jcim.0c00464.
22. Long, T., Fortunato, N.M., Opahle, I., Zhang, Y., Samathrakis, I., Shen, C., Gutfleisch, O., and Zhang, H. (2021). Constrained crystals deep convolutional generative adversarial network for the inverse design of crystal structures. npj Comput. Mater. *7*, 66. https://doi.org/10.1038/s41524-021-00526-4.
23. Zhao, Y., Al-Fahdi, M., Hu, M., Siriwardane, E.M.D., Song, Y., Nasiri, A., and Hu, J. (2021). High-Throughput Discovery of Novel Cubic Crystal Materials Using Deep Generative Neural Networks. Adv. Sci. *8*, 2100566. https://doi.org/10.1002/advs.202100566.
24. Ren, Z., Tian, S.I.P., Noh, J., Oviedo, F., Xing, G., Li, J., Liang, Q., Zhu, R., Aberle, A.G., Sun, S., et al. (2022). An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. Matter *5*, 314–335. https://doi.org/10.1016/j.matt.2021.11.032.
25. Xie, T., Fu, X., Ganea, O.-E., Barzilay, R., and Jaakkola, T. (2022). Crystal Diffusion Variational Autoencoder for Periodic Material Generation. Preprint at arXiv. https://doi.org/10.48550/arXiv.2110.06197.
26. Antunes, L.M., Butler, K.T., and Grau-Crespo, R. (2023). Crystal Structure Generation with Autoregressive Large Language Modeling. Preprint at arXiv. https://doi.org/10.48550/arXiv.2307.04340.
27. Flam-Shepherd, D. (2023). Aspuru-Guzik A. Language models can generate molecules, materials, and protein binding sites directly in three dimensions as XYZ, CIF, and PDB files.

Preprint at arXiv. https://doi.org/10.48550/arXiv.2305.05708.

28. Alverson, M., Baird, S., Murdock, R., Ho, E.S.-H., Johnson, J., and Sparks, T. (2024). Generative adversarial networks and diffusion models in material discovery. Digital Discovery *3*, 62–80. https://doi.org/10.1039/D3DD00137G.

29. Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y., and Liu, Y. (2024). Crystal Structure Prediction by Joint Equivariant Diffusion. Preprint at arXiv. [cond-mat]. https://doi.org/10.48550/arXiv.2309.04475.

30. Zeni, C., Pinsler, R., Zügner, D., Fowler, A., Horton, M., Fu, X., Shysheya, S., Crabbé, J., Sun, L., Smith, J., et al. (2023). MatterGen: a generative model for inorganic materials design. Preprint at arXiv. https://doi.org/10.48550/arXiv.2312.03687.

31. Zhu, R., Nong, W., Yamazaki, S., and Hippalgaonkar, K. (2023). WyCryst: Wyckoff Inorganic Crystal Generator Framework. Preprint at arXiv. https://doi.org/10.48550/arXiv.2311.17916.

32. Yang, M., Cho, K., Merchant, A., Abbeel, P., Schuurmans, D., Mordatch, I., and Cubuk, E.D. (2023). Scalable Diffusion for Materials Generation. Preprint at arXiv. https://doi.org/10.48550/arXiv.2311.09235.

33. Gruver, N., Sriram, A., Madotto, A., Wilson, A.G., Zitnick, C.L., and Ulissi, Z. (2024). Fine-tuned language models generate stable inorganic materials as text. Preprint at arXiv. https://doi.org/10.48550/arXiv.2402.04379.

34. Musil, F., Grisafi, A., Bartók, A.P., Ortner, C., Csányi, G., and Ceriotti, M. (2021). Physics-Inspired Structural Representations for Molecules and Materials. Chem. Rev. *121*, 9759–9815. publisher: American Chemical Society. https://doi.org/10.1021/acs.chemrev.1c00021.

35. Türk, H., Landini, E., Kunkel, C., Margraf, J.T., and Reuter, K. (2022). Assessing Deep Generative Models in Chemical Composition Space. Chem. Mater. *34*, 9455–9467. publisher: American Chemical Society. https://doi.org/10.1021/acs.chemmater.2c01860.

36. Dai, B., Wang, Z., and Wipf, D. (2020). The Usual Suspects? Reassessing Blame for VAE Posterior Collapse. In Proceedings of the 37th International Conference on Machine Learning (PMLR), pp. 2313–2322. iSSN: 2640-3498.

37. Dhariwal, P., and Nichol, A. (2021). Diffusion models beat gans on image synthesis. Preprint at arXiv. https://doi.org/10.48550/arXiv.2105.05233.

38. Ho, J., Jain, A., and Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. Preprint at arXiv. https://doi.org/10.48550/arXiv.2006.11239.

39. Ambrogioni, L. (2024). The statistical thermodynamics of generative diffusion models: Phase transitions, symmetry breaking and critical instability. Preprint at arXiv. [cs, stat]. https://doi.org/10.48550/arXiv.2310.17467.

40. Song, Y., and Ermon, S. (2020). Generative Modeling by Estimating Gradients of the Data Distribution. Preprint at arXiv. [cs, stat]. https://doi.org/10.48550/arXiv.1907.05600.

41. Ho, J., and Salimans, T. (2022). Classifier-free diffusion guidance. Preprint at. arXiv. https://doi.org/10.48550/arXiv.2207.12598.

42. Park, J., Gill, A.P.S., Moosavi, S.M., and Kim, J. (2024). Inverse design of porous materials: a diffusion model approach. J. Mater. Chem. A *12*, 6507–6514. https://doi.org/10.1039/D3TA06274K.

43. Dagdelen, J., Dunn, A., Lee, S., Walker, N., Rosen, A.S., Ceder, G., Persson, K.A., and Jain, A. (2024). Structured information extraction from scientific text with large language models. Nat. Commun. *15*, 1418. publisher: Nature Publishing Group. https://doi.org/10.1038/s41467-024-45563-x.

44. Jablonka, K.M., Schwaller, P., Ortega-Guerrero, A., and Smit, B. (2024). Leveraging large language models for predictive chemistry. Nat. Mach. Intell. *6*, 161–169. https://doi.org/10.1038/s42256-023-00788-1.

45. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2023). Attention Is All You Need. Preprint at arXiv. [cs]. https://doi.org/10.48550/arXiv.1706.03762.

46. Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., et al. (2023). Llama 2: Open foundation and fine-tuned chat models. Preprint at arXiv. https://doi.org/10.48550/arXiv.2307.09288.

47. Leeman, J., Liu, Y., Stiles, J., Lee, S.B., Bhatt, P., Schoop, L.M., and Palgrave, R.G. (2024). Challenges in High-Throughput Inorganic Materials Prediction and Autonomous Synthesis. PRX Energy *3*, 011002. https://doi.org/10.1103/PRXEnergy.3.011002.

48. Cheetham, A.K., and Seshadri, R. (2024). Artificial Intelligence Driving Materials Discovery? Perspective on the Article: Scaling Deep Learning for Materials Discovery. Chem. Mater. *36*, 3490–3495. https://doi.org/10.1021/acs.chemmater.4c00643.

49. Davies, D.W., Butler, K.T., Isayev, O., and Walsh, A. (2018). Materials discovery by chemical analogy: role of oxidation states in structure prediction. Faraday Discuss *211*, 553–568. https://doi.org/10.1039/C8FD00032H.

50. Batatia, I., Benner, P., Chiang, Y., Elena, A.M., Kovács, D.P., Riebesell, J., Advincula, X.R., Asta, M., Baldwin, W.J., Bernstein, N., et al. (2023). A foundation model for atomistic materials chemistry. Preprint at arXiv. https://doi.org/10.48550/arXiv.2401.00096.

51. Brown, I.D. (1969). Crystal structures of $NaNiIO_6$, $NaMnIO_6$, and $KMnIO_6$, Canad. J. Chem. *47*, 3779–3782.

52. Duan, L., Zhou, T., Zhang, Y., Zhao, J., Zheng, H., Zi, B., Zhang, J., Li, Q., Liu, J., and Liu, Q. (2023). Surface optics and color effects of liquid metal materials. Adv. Mater. *35*, 2210515.

53. Benny, Y., Galanti, T., Benaim, S., and Wolf, L. (2021). Evaluation metrics for conditional image generation. Int. J. Comput. Vis. *129*, 1712–1731. https://doi.org/10.1007/s11263-020-01424-w.

54. Ong, S.P., Richards, W.D., Jain, A., Hautier, G., Kocher, M., Cholia, S., Gunter, D., Chevrier, V.L., Persson, K.A., and Ceder, G. (2013). Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. Comput. Mater. Sci. *68*, 314–319. https://doi.org/10.1016/j.commatsci.2012.10.028.

55. matbench-genmetrics (2024). A Collection of Metrics for Benchmarking Materials Science and Chemistry Models. https://github.com/sparks-baird/matbench-genmetrics.

56. Schön, J.C., and Jansen, M. (1996). First Step Towards Planning of Syntheses in Solid-State Chemistry: Determination of Promising Structure Candidates by Global Optimization. Angew. Chem., Int. Ed. Engl. *35*, 1286–1304. https://doi.org/10.1002/anie.199612861.

57. Tolborg, K., Klarbring, J., Ganose, A.M., and Walsh, A. (2022). Free energy predictions for crystal stability and synthesisability. Digital Discovery *1*, 586–595. https://doi.org/10.1039/D2DD00050D.

58. Reilly, A.M., Cooper, R.I., Adjiman, C.S., Bhattacharya, S., Boese, A.D., Brandenburg, J.G., Bygrave, P.J., Bylsma, R., Campbell, J.E., Car, R., et al. (2016). Report on the sixth blind test of organic crystal structure prediction methods. Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater. *72*, 439–459. https://doi.org/10.1107/S2052520616007447.

59. Rønne, N., Aspuru-Guzik, A., Hammer, B. (2024) Generative Diffusion Model for Surface Structure Discovery. Preprint at arXiv. https://doi.org/10.48550/arxiv.2402.17404.

60. Li, Z., Nash, W., O'Brien, S., Qiu, Y., Gupta, R., and Birbilis, N. (2022). cardiGAN: A generative adversarial network model for design and discovery of multi principal element alloys. J. Mater. Sci. Technol. *125*, 81–96. https://doi.org/10.1016/j.jmst.2022.03.008.

61. Zhou, Z., Shang, Y., Liu, X., and Yang, Y. (2023). A generative deep learning framework for inverse design of compositionally complex bulk metallic glasses. npj Comput. Mater. *9*, 15. https://doi.org/10.1038/s41524-023-00968-y.

62. Banko, L., Lysogorskiy, Y., Grochla, D., Naujoks, D., Drautz, R., and Ludwig, A. (2020). Predicting structure zone diagrams for thin film synthesis by generative machine learning. Commun. Mater. *1*, 15. https://doi.org/10.1038/s43246-020-0017-2.

63. Gómez-Bombarelli, R., Wei, J.N., Duvenaud, D., Hernández-Lobato, J.M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T.D., Adams, R.P., and Aspuru-Guzik, A. (2018). Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. ACS Cent. Sci. *4*, 268–276. https://doi.org/10.1021/acscentsci.7b00572.