International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Artificial Intelligence in the Discovery of Functional Solid-State Materials: Opportunities, Challenges, and Future Perspectives

Narendra Singh Lakhera

Dr. Bhimrao Ambedkar Government P.G. College Mahwa (Dausa)

Abstract

Data-rich, closed-loop enterprise is transforming functional solid-state materials discovery from a sequential process of intuition, synthesis, characterization, and delayed theory to an artificial intelligence (AI) enabled process. This review discusses the application of machine learning, graph neural networks, active learning, high-throughput density functional theory (DFT), universal machine-learned interatomic potentials and autonomous laboratories for discovering inorganic crystals, solid electrolytes, semiconductors, magnets, thermoelectric and other functional materials. The paper claims that the most promising opportunity is not merely to replace physics-based modelling, but to couple AI with DFT, thermodynamics, synthesis knowledge and uncertainty-aware experimental validation. Reported demonstrations, such as GNoME-scale screening, graph-based property prediction, charge-informed potentials, and autonomous powder synthesis show that AI can increase candidate spaces by orders of magnitude and decrease wasted experiments. Yet, data bias, failure to report failed experiments, synthesis complexity, benchmark leakage, out-of-distribution errors, weak interpretability, and lack of negative-data reporting still remain substantial barriers to SCI-level reproducibility. Future advances will rely on FAIR data infrastructures, prospective benchmarks, hybrid physics-AI models, human-auditable autonomous laboratories and application-specific validation pipelines. The discovery of functional solid-state materials should therefore be considered as an integrated cyber-physical system and not merely as a computational ranking exercise.

Keywords: Artificial intelligence; Functional materials; Solid-state materials; Graph neural networks; Materials informatics; Autonomous laboratory; Machine-learned interatomic potentials

1. Introduction

Functional solid-state materials are crystalline inorganic compounds that have electronic, ionic, optical, magnetic, catalytic, thermal, or mechanical responses that enable technologies such as batteries, photovoltaics, sensors, information devices, membranes, and quantum hardware. Their discovery has been traditionally dependent on expert chemical intuition, incremental substitutions, trial-and-error synthesis and expensive characterization. This model is scientifically productive, but slow, because the relevant design space is combinatorial: composition, crystal structure, defects, microstructure, processing route and operating environment all interact nonlinearly. AI has become attractive in that it is able to search a large chemical space, learn structure-property relationships from accumulated data and can steer experiments to regions of higher expected information gain (Butler et al., 2018; Ramprasad et al., 2017).

The field has seen progress from early descriptor-based regression to graph neural networks, active learning and universal machine learned interatomic potentials. The Materials Genome Initiative set the broader paradigm that theory, computation, data and experiment must be integrated to accelerate materials deployment (de Pablo et al., 2019). In this context, AI is best viewed as an enabling layer that translates heterogeneous knowledge into decisions: which compositions to compute, which structures to relax, which synthesis recipes to test, which diffraction patterns are convincing and which uncertainties should be reviewed by a human. So, AI alone is not enough to discover functional solids, but AI combined with physics is.

In this paper, we review the opportunities, challenges and future perspectives in AI-driven discovery of functional solid-state materials in a SCI style. It focuses on inorganic crystalline systems, which are the driving force of recent advances in graph learning, high-throughput DFT, and autonomous solid-state synthesis. The crux of the matter is that AI can only revolutionize materials discovery when its predictions are linked to thermodynamic stability, synthesizability, device-relevant functionality, and reproducible validation.

The measure of success changes too, with a functional-materials perspective. A candidate just below the convex hull may still fail as a technology if it contains scarce elements, decomposes at an interface, requires impractical synthesis conditions or loses performance after cycling. Conversely, a metastable phase may be of technological utility if it can be reproducibly synthesized and retained under service conditions. Thus, AI discovery should prioritize materials by a hierarchy of evidence: calculated stability, feasible synthesis, measurable property, durability, and system-level usefulness.

2. Scope and Review Approach

The review summarizes the peer-reviewed work on data infrastructures, representation learning, graph neural networks, interatomic potentials, active learning and autonomous laboratories. The emphasis is on functional properties and not structural materials per se, although mechanical and thermal descriptors are included where they affect device performance. Studies were prioritized if they offered a reusable data source, a broadly applicable model architecture, a quantitative discovery demonstration, or a clear critique of current AI practices in materials science.

The review is narrative instead of meta-analytic, since benchmark datasets, target properties and validation methods differ significantly in the literature. The formation energy, the band gap, the decomposition energy, the ionic conductivity and the synthesis success cannot be combined into a single universal effect size without losing physical meaning. Instead, the paper compares model classes and workflows by task, scale, validation mode and known limitations. This is a typical way that review papers in fast moving interdisciplinary fields organize their material in ScienceDirect: first describe data and methods, and then discuss opportunities, challenges and directions for future research.

The term AI is used broadly, but not loosely. This paper covers supervised property prediction, deep representation learning, graph neural networks, active learning, generative or enumerative candidate construction, machine-learned interatomic potentials, and decision policies for autonomous laboratories. Pure database search is not thought of as AI unless there is a learning algorithm or an adaptive decision rule. This boundary is important because many inflated claims are made when routine high-throughput screening is rebranded as artificial intelligence, with no predictive or adaptive component.

Element	Selected value	Rationale
Article category	SCI-style review paper	Narrative review with quantitative literature tables; no new experiments performed.
Domain focus	Functional solid-state materials	Inorganic crystals, solid electrolytes, semiconductors, magnets, thermoelectrics, dielectric and catalytic solids.
AI scope	ML, deep learning, GNNs, active learning, MLIPs, autonomous labs	Large language models are considered only where linked to structured synthesis workflows.
Evidence priority	Peer-reviewed articles	Databases, model papers, benchmarks and autonomous synthesis demonstrations.
Main evaluation lens	Opportunity vs. reproducibility risk	Predicted stability alone is treated as insufficient for application readiness.

Table 1: Review scope and inclusion logic, Source: Author Generated

3. Data Foundations for AI Materials Discovery

Structured data linking composition and crystal structure to properties are required for AI models of solid-state materials. The first big opportunity was from the high throughput DFT databases. The Materials Project, AFLOWLIB, OQMD and NOMAD turned individual calculations into searchable infrastructures where researchers can train models for formation energies, band gaps, elastic constants and stability indicators (Curtarolo et al., 2012; Draxl & Scheffler, 2019; Jain et al., 2013; Kirklin et al., 2015). These platforms have reduced the cost of candidate screening by enabling a model to learn from thousands to millions of reference calculations before attempting expensive experiments.

But availability of data is not the same as appropriateness of data. Systematic errors in DFT labels stem from functionals, pseudopotentials and convergence settings; experimental data suffer from impurities, processing history and inconsistent reporting. Survivorship bias is a result of the usual failure to deposit failed syntheses. Target-specific data is also crucial for finding functional materials: a stable crystal with low ion mobility, optical absorption or defect tolerance is not useful. Therefore, AI pipelines need to record provenance, uncertainty, and missingness, rather than treating each database entry as an equally reliable fact.

The data problem for functional solids is multi-dimensional. A solid electrolyte for battery requires phase stability, electrochemical window, lithium or sodium mobility, grain-boundary behavior and electrode compatibility. A band gap, effective masses, defect tolerance, dupability and optical absorption are required for a semiconductor. A thermoelectric needs electronic transport, phonons, mechanical stability and processing tolerance. No one database captures all of these properties with the same level of reliability, so the next generation of AI discovery will depend on data fusion across computations, experiments, text mining and laboratory automation.

Platform	Type	Quantitative or functional data point	Representative citation
Materials Project	High-throughput DFT database and analysis ecosystem	Computed structures and properties for known and predicted materials	Jain et al. (2013)
AFLOWLIB	Automated high-throughput ab initio repository	Reported >150,000 thermodynamic alloy entries in early release context	Curtarolo et al. (2012)
OQMD	Open Quantum Materials Database	Compared DFT formation energies with 1,670 experimental formation energies	Kirklin et al. (2015)
NOMAD	FAIR materials-data platform	Code-independent normalization and AI-ready data sharing	Draxl & Scheffler (2019)
GNoME dataset	Active-learning discovery catalogue	2.2 million stable structures versus prior work; 381,000 on updated convex hull	Merchant et al. (2023)

Table 2: Major data infrastructures enabling AI-driven solid-state materials discovery, Source: Author Generated

4. Model Families and Opportunities

The early materials informatics was based on composition descriptors and kernel or tree models. These methods work even if structures are unknown, which is often the case at the beginning of discovery. Ward et al. (2016) demonstrated that a set of reusable chemical attributes could be used to predict a broad spectrum of inorganic properties and ElemNet demonstrated that deep learning could be used to predict chemical trends based solely on elemental composition (Jha et al., 2018). Goodall and Lee (2020) further extended the structure-agnostic learning to include stoichiometry as a dense graph between elements. The opportunity for the practitioner, then, is that AI can triage compositions for likely stability or functionality even before a crystal structure has been resolved.

Structure-aware Graph Neural Networks are a subset of a larger methodological shift. The CGCNN encodes atoms and bonds in crystals such that properties are learned from local environments, rather than from hand-engineered fingerprints (Xie & Grossman, 2018). MEGNet (Chen et al., 2019) generalised this idea for molecules and crystals with graph networks. ALIGNN (Choudhary & DeCost, 2021) introduced line-graph message passing to explicitly consider bond-angle information. These models are particularly important for functional solids, since band structure, ion migration, magnetism and phonons are all functions of geometry and composition.

The latest possibility is atomistic modelling in foundation style. Graph models trained on large relaxation datasets can approximate potential-energy surfaces across wide chemical domains, as demonstrated by M3GNet and CHGNet (Chen & Ong, 2022; Deng et al., 2023). This scale was pushed further by GNoME using active learning and graph networks to search millions of predicted stable crystals, including hundreds of thousands on a modified convex hull (Merchant et al., 2023). These models can reduce the number of DFT calculations needed for each discovery campaign and allow downstream simulations that would be prohibitive otherwise.

Active learning is the operational bridge between prediction and discovery. Active-learning workflows select candidates iteratively to maximise expected improvement, reduce uncertainty or test model assumptions, rather than training a model once and screening a fixed library. Meredig et al., 2014; Pilania et al., 2013) and helped establish machine learning as a useful tool to prioritise new material spaces before exhaustive calculations or synthesis. In modern graph-based active learning the same logic is applied to crystal structures, where each new DFT relaxation can be used to validate a candidate and to improve the next model. The main opportunity is efficiency. In scientific terms, the value of a calculation goes up if it is chosen to be informative, rather than simply available.

Model family	Core idea	Typical task	Representative citations
Composition descriptors	Uses elemental attributes; useful when structures are unknown	Band gap, glass-forming ability, stability triage	Ward et al. (2016); Meredig et al. (2014)
Deep composition networks	Learns chemistry directly from stoichiometry	Formation energy or other property screening	Jha et al. (2018); Goodall & Lee (2020)
Crystal graph neural networks	Atoms and bonds encoded as graph nodes and edges	Structure-property learning across crystals	Xie & Grossman (2018); Chen et al. (2019)
Line-graph GNNs	Adds bond-angle message passing	52 solid-state and molecular prediction tasks reported	Choudhary & DeCost (2021)
Universal MLIPs	Predict energies/forces for broad chemistries	Fast relaxation, dynamics, diffusion and thermodynamics	Chen & Ong (2022); Deng et al. (2023)
Active-learning discovery	Model proposes candidates, DFT validates and retrains	Large-scale stable-crystal discovery	Merchant et al. (2023)

Table 3: AI model families and discovery roles in functional solid-state materials, Source: Author Generated

Example	Discovery or prediction task	Reported data point	Why it matters	Citation
CGCNN	Crystal property prediction	Trained on about 10^4 DFT data points across eight properties	Interpretable graph representation for crystals	Xie & Grossman (2018)
MEGNet	Molecules and crystals	Outperformed prior models on 11 of 13 QM9 properties in reported comparison	Unified graph framework	Chen et al. (2019)
M3GNet	Universal potential	Screened 31 million hypothetical structures; 1.8 million predicted potentially stable	Fast broad chemistry relaxation	Chen & Ong (2022)
CHGNet	Charge-informed MLIP	Pretrained on >1.5 million inorganic structures from MP trajectories	Includes magnetic moment information	Deng et al. (2023)
GNoME	Stable inorganic crystals	381,000 new convex-hull entries; 736 independently realized	Order-of-magnitude expansion of known stable crystals	Merchant et al. (2023)
A-Lab	Autonomous synthesis	Reported continuous 17-day operation with robotic synthesis and ML-guided recipe updates	Closed-loop experimental validation	Szymanski et al. (2023)

5. Closed-Loop Discovery and Autonomous Laboratories

The strongest vision for AI in functional solid-state materials is a closed loop in which the model proposes candidates, computation verifies stability, robotics performs synthesis, characterization yields data, and active learning selects the next experiment. Correa-Baena et al. (2018) posited that automation, machine learning and high-performance computing could shorten the development timelines of materials by at least an order of magnitude in favorable cases. This concept here was translated to mainly inorganic powder synthesis with the A-Lab demonstration combining here after computed phase stability, literature-derived synthesis recipes, robotic experimentation, X-ray diffraction analysis & active learning (Szymanski et al., 2023).

Closed-loop systems are powerful because they learn from failure, rather than hide it. Even if a target fails to form, the recipe, precursor choice, heating profile and phase analysis still provide useful information. This is particularly true in solid state chemistry where kinetic barriers, competing phases, air sensitivity and precursor availability often determine success over thermodynamic stability alone. In this way AI can help to discover functional materials by optimizing the synthesis conditions, not only by ranking idealized crystals. But autonomous laboratories have to report validation criteria transparently, as a seeming success can be dependent on phase-purity thresholds, diffraction refinement quality and whether the product is really new.

A closed loop also alters the function of the scientist. For human researchers, manual assessment of every low-value candidate is no longer necessary, but the definition of objectives, safety limits, acceptable uncertainty and stopping criteria is still needed. For instance, an AI system might optimize the yield of a predicted phase, but overlook impurity phases that hinder device performance. It can also take advantage of biases in an automated phase identification model. Thus, human oversight is not a weakness of autonomous discovery, but the mechanism that keeps optimization aligned with scientific meaning and application needs.

6. Challenges

1. **The first challenge** is domain shift. Materials discovery is explicitly extrapolative, i.e. the goal is to discover compounds outside the training distribution. It may have low average error on retrospective benchmarks but poor performance on rare chemistries, high energy defect states or metastable phases. To partly address this issue, Matbench Discovery was developed to assess models as pre-filters to stability searches rather than static predictors only (Riebesell et al., 2025). The risk is also shown by universal potentials. Deng et al. (2025) showed systematic potential-energy-surface softening in several universal MLIPs, especially in out-of-distribution high-energy configurations. This is important for diffusion barriers, defects, surfaces and reactions that are central to functionality.
2. **The second challenge** is interpretability. Graph neural networks are capable of discovering chemically meaningful patterns but their latent representations are not explanations by default. A model might learn correlations with database conventions, rather than causal physics. For example, predictions of band gaps might be numerically better but still inherit the DFT underestimation, and stability rankings might be biased toward known prototypes. AI predictions for SCI-level research should provide uncertainty estimates, ablation tests, external validation, and, when feasible, physically interpretable descriptors.
3. **The third challenge** is replicability Many papers on AI materials rely on evolving databases, private code, undocumented preprocessing or train-test splits that inadvertently leak structural families. Reproducibility requires versioned datasets, published model weights, explicit definitions of features, and prospective tests on materials unknown at training time. The field needs more experimental data that are negative. If some syntheses do not work, the AI systems learn only what has been reported, not what is possible.
4. **The fourth challenge** is evaluation at the boundary from property-to-device. Many manuscripts report a predicted band gap, elastic constant or ionic conductivity as if it is a material solution. In practice the functional performance is determined by interfaces, defects, processing, stability under fields or temperature and manufacturability. This is why AI-discovered solid-state materials must be screened with application-aware funnels. For example, a candidate solid electrolyte must go through thermodynamic, kinetic, mechanical and electrochemical filters to be considered battery-relevant. This more conservative language will reduce hype and increase translational value.

Challenge	Mechanism	Risk for discovery	Mitigation
Data bias	Known structures and successful syntheses are over-represented	Survivorship bias and false confidence	Publish failed experiments and provenance metadata
DFT label error	Functionals and settings introduce systematic deviations	Biased band gaps, stability and reaction energetics	Use correction schemes and multi-fidelity validation
Domain shift	Discovery targets are outside training chemistry	Retrospective error metrics overstate usefulness	Prospective benchmarks and uncertainty calibration
Synthesis gap	Thermodynamic stability does not ensure a feasible route	Predicted materials remain hypothetical	Couple AI with synthesis models, kinetics and robotics
Interpretability	Latent graph features may not map to mechanisms	Low trust and weak scientific explanation	Use ablations, descriptors, SHAP-like tools and physics constraints
Reproducibility	Data versions and preprocessing are often unclear	Difficult replication and comparison	Release code, splits, weights and database snapshots

Table 4: Major challenges and mitigation strategies, Source: Author Generated

7. Future Perspectives

The future of AI-driven discovery will shift from model-centric papers to workflow-centric platforms. The most useful systems will combine composition models for early triage, GNNs for structure-property prediction, universal MLIPs for fast relaxation and dynamics, thermodynamic filters for stability, synthesis models for route design and robotic feedback for experimental validation. Large language models can be used for literature mining and recipe drafting, but they have to be anchored to structured databases and audited decision rules.

Functional performance must also be aimed at, not added on at the end. Many discovery campaigns stop at predicted stability, but device materials need multiple constraints: processability, abundance, toxicity, durability, interfaces, defect chemistry and performance under operating conditions. AI can assist with multi-objective optimisation, but only if the objectives are well defined and validated. For solid electrolytes, this involves a compromise between stability, ionic conductivity and electrochemical compatibility and, for photovoltaics, the band gap, the absorption, defect tolerance and manufacturability have to be considered together.

A second future direction is physics-informed and uncertainty-aware learning. Models should encode symmetry and conservation laws, charge states and thermodynamic constraints, reporting calibrated uncertainty. Active learning should focus on promising and informative candidates. Finally, community benchmarks should become prospective discovery challenges in which predictions are made prior to new calculations or experiments. This would separate genuine discovery skill from post hoc curve fitting and would make AI materials science more credible to industrial and academic users.

Future papers should also report environmental and supply-chain constraints earlier in the discovery process. Scientifically interesting materials may also contain rare, toxic or geopolitically constrained elements, rendering them unattractive for scalable technologies. Incorporating abundance, price volatility, recyclability and toxicity into multi-objective AI models would foster more responsible and relevant discoveries for the clean-energy and electronics industries. These socio-technical filters are not meant to replace fundamental research, but they should be visible when claims are made about future applications.

Time horizon	Priority	Required technical advance	Expected impact
2026-2028	Benchmark discipline	Prospective discovery tasks, versioned datasets, uncertainty-aware leaderboards	Fewer inflated claims; stronger model comparison
2028-2030	Hybrid physics-AI workflows	Thermodynamics, symmetry, charge and kinetics integrated into learning	Better transfer to defects, interfaces and metastable phases
2030-2032	Autonomous materials platforms	Robotic synthesis and characterization linked to active learning	Routine closed-loop optimization for selected material classes
2032 onward	Application-ready discovery	Multi-objective optimization includes cost, toxicity, durability and device metrics	Faster transition from prediction to deployable material

Table 5: Future roadmap for AI-enabled functional solid-state materials discovery, Source: Author Generated

8. Conclusion

AI is building a new infrastructure for discovering functional solid-state materials. It can mine databases, learn composition-structure-property relationships, accelerate DFT workflows, propose stable crystals, approximate interatomic potentials and guide autonomous laboratories. Graph-based discovery campaigns and self-driving synthesis platforms show there is a huge opportunity. But evidence suggests AI is not a substitute for chemistry. Stable predictions require routes to their synthesis, functional validation, and repeatable evidence. Therefore, the next step in the field should be the focus on integrated, interpretable and uncertainty-aware cyber-physical workflows. Coupled with rigorous thermodynamics, open data, robotic experimentation and human chemical judgement, AI can shorten the path from hypothetical crystal to useful functional material.

References

1. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. Nature, 559(7715), 547-555. https://doi.org/10.1038/s41586-018-0337-2
2. Chen, C., & Ong, S. P. (2022). A universal graph deep learning interatomic potential for the periodic table. Nature Computational Science, 2, 718-728. https://doi.org/10.1038/s43588-022-00349-3

3. Chen, C., Ye, W., Zuo, Y., Zheng, C., & Ong, S. P. (2019). Graph networks as a universal machine learning framework for molecules and crystals. Chemistry of Materials, 31(9), 3564-3572. https://doi.org/10.1021/acs.chemmater.9b01294
4. Choudhary, K., & DeCost, B. (2021). Atomistic line graph neural network for improved materials property predictions. npj Computational Materials, 7, 185. https://doi.org/10.1038/s41524-021-00650-1
5. Correa-Baena, J. P., Hippalgaonkar, K., van Duren, J., Jaffer, S., Chandrasekhar, V. R., Stevanovic, V., Wadia, C., Guha, S., & Buonassisi, T. (2018). Accelerating materials development via automation, machine learning, and high-performance computing. Joule, 2(8), 1410-1420. https://doi.org/10.1016/j.joule.2018.05.009
6. Curtarolo, S., Setyawan, W., Wang, S., Xue, J., Yang, K., Taylor, R. H., Nelson, L. J., Hart, G. L. W., Sanvito, S., Buongiorno-Nardelli, M., Mingo, N., & Levy, O. (2012). AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. Computational Materials Science, 58, 227-235. https://doi.org/10.1016/j.commatsci.2012.02.002
7. de Pablo, J. J., Jackson, N. E., Webb, M. A., Chen, L. Q., Moore, J. E., Morgan, D., Jacobs, R., Pollock, T., Schlom, D. G., Toberer, E. S., Analytis, J., Dabo, I., DeLongchamp, D. M., Fiete, G. A., Grason, G. M., Hautier, G., Mo, Y., Rajan, K., Reed, E. J., Rodriguez, E., Stevanovic, V., Suntivich, J., Thornton, K., & Zhao, J. C. (2019). New frontiers for the materials genome initiative. npj Computational Materials, 5, 41. https://doi.org/10.1038/s41524-019-0173-4
8. Deng, B., Choi, Y., Zhong, P., Riebesell, J., Anand, S., Li, Z., Jun, K., Persson, K. A., & Ceder, G. (2025). Systematic softening in universal machine learning interatomic potentials. npj Computational Materials, 11, 9. https://doi.org/10.1038/s41524-024-01500-6
9. Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C. J., & Ceder, G. (2023). CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. Nature Machine Intelligence, 5, 1031-1041. https://doi.org/10.1038/s42256-023-00716-3
10. Draxl, C., & Scheffler, M. (2019). The NOMAD laboratory: From data sharing to artificial intelligence. Journal of Physics: Materials, 2(3), 036001. https://doi.org/10.1088/2515-7639/ab13bb
11. Goodall, R. E. A., & Lee, A. A. (2020). Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. Nature Communications, 11, 6280. https://doi.org/10.1038/s41467-020-19964-7
12. Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., & Persson, K. A. (2013). Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. APL Materials, 1(1), 011002. https://doi.org/10.1063/1.4812323
13. Jha, D., Ward, L., Paul, A., Liao, W. K., Choudhary, A., Wolverton, C., & Agrawal, A. (2018). ElemNet: Deep learning the chemistry of materials from only elemental composition. Scientific Reports, 8, 17593. https://doi.org/10.1038/s41598-018-35934-y
14. Kirklin, S., Saal, J. E., Meredig, B., Thompson, A., Doak, J. W., Aykol, M., Ruhl, S., & Wolverton, C. (2015). The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies. npj Computational Materials, 1, 15010. https://doi.org/10.1038/npjcompumats.2015.10

15. Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. Nature, 624(7990), 80-85. https://doi.org/10.1038/s41586-023-06735-9
16. Meredig, B., Agrawal, A., Kirklin, S., Saal, J. E., Doak, J. W., Thompson, A., Zhang, K., Choudhary, A., & Wolverton, C. (2014). Combinatorial screening for new materials in unconstrained composition space with machine learning. Physical Review B, 89(9), 094104. https://doi.org/10.1103/PhysRevB.89.094104
17. Pilania, G., Wang, C., Jiang, X., Rajasekaran, S., & Ramprasad, R. (2013). Accelerating materials property predictions using machine learning. Scientific Reports, 3, 2810. https://doi.org/10.1038/srep02810
18. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., & Kim, C. (2017). Machine learning in materials informatics: Recent applications and prospects. npj Computational Materials, 3, 54. https://doi.org/10.1038/s41524-017-0056-5
19. Riebesell, J., Goodall, R. E. A., Jain, A., Benner, P., Persson, K. A., & Lee, A. A. (2025). A framework to evaluate machine learning crystal stability predictions. Nature Machine Intelligence, 7, 809-819. https://doi.org/10.1038/s42256-025-01055-1
20. Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials, 5, 83. https://doi.org/10.1038/s41524-019-0221-0
21. Szymanski, N. J., Rendy, B., Fei, Y., Kumar, R. E., He, T., Milsted, D., McDermott, M. J., Gallant, M., Cubuk, E. D., Merchant, A., Kim, H., Jain, A., Bartel, C. J., Persson, K., Zeng, Y., & Ceder, G. (2023). An autonomous laboratory for the accelerated synthesis of novel materials. Nature, 624(7990), 86-91. https://doi.org/10.1038/s41586-023-06734-w
22. Ward, L., Agrawal, A., Choudhary, A., & Wolverton, C. (2016). A general-purpose machine learning framework for predicting properties of inorganic materials. npj Computational Materials, 2, 16028. https://doi.org/10.1038/npjcompumats.2016.28
23. Xie, T., & Grossman, J. C. (2018). Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. Physical Review Letters, 120(14), 145301. https://doi.org/10.1103/PhysRevLett.120.145301