

REVIEW

Open access

Representation Learning in Materials Science: Architectures, Data Modalities, and Discovery Applications

Claire Martin^{1*}, Julien Martin², Sophie Bernard¹

Abstract

The field of materials science has undergone a transformative shift with the integration of computational and data-driven approaches, particularly through representation learning techniques that enable efficient handling of complex materials data. This review synthesizes recent advancements in architectures for representation learning, encompassing graph neural networks, attention-based models, and physics-inspired embeddings, which facilitate the extraction of meaningful features from diverse data modalities such as atomic structures, stoichiometries, and spectroscopic data. By bridging traditional computational methods with machine learning, these representations have accelerated property prediction, inverse design, and materials discovery applications, addressing challenges in high-dimensional spaces and sparse datasets. The scope of this narrative review covers the evolution from basic informatics to sophisticated multimodal integrations, highlighting how data ecosystems and learning frameworks contribute to autonomous discovery pipelines. A systems-level perspective is adopted to integrate cross-study insights, revealing synergies between representation learning and closed-loop systems that couple simulations with experiments. Looking ahead, the review posits that continued refinement of these architectures will drive scalable, AI-guided materials engineering, fostering innovations in energy, electronics, and structural materials while emphasizing the need for robust, interpretable models in real-world applications.

Keywords Materials informatics, Property prediction, Inverse design, Graph neural networks, Representation learning, Data modalities

*Correspondence:

Claire Martin

claire.martin@gmail.com

¹ Department of Computational Materials Systems, Faculty of Engineering, University of Lyon, Lyon, France

² Department of Materials Data Analytics, Faculty of Engineering, University of Strasbourg, Strasbourg, France

Introduction

Materials science has historically relied on empirical experimentation and theoretical modeling to uncover new compounds and optimize properties, but the advent of computational tools has dramatically expanded the scope and speed of discovery. The evolution of computational materials science began with density functional theory (DFT) simulations in the mid-20th century, enabling atomic-scale predictions of electronic structures and energies [1]. However, these methods were computationally intensive, limiting their application to small systems. The integration of high-throughput computing in the early 2000s marked a pivotal shift, allowing for the systematic screening of vast materials spaces through databases like the Materials Project and Automatic FLOW for Materials Discovery (AFLOW) [2]. This era emphasized brute-force enumeration, yet it highlighted the need for smarter, data-centric strategies to navigate the immense combinatorial complexity of materials compositions and structures.

The rise of AI and data-driven infrastructures has further revolutionized the field, transitioning from passive simulation to active, predictive modeling. Machine learning (ML) techniques, initially applied to regression tasks for property prediction, have evolved into sophisticated representation learning paradigms that encode materials information in low-dimensional, informative vectors [3, 4]. These representations capture essential physicochemical features, such as atomic neighborhoods, symmetries, and electronic configurations, enabling generalization across diverse materials classes [5, 6]. For instance, stoichiometry-based embeddings have demonstrated the ability to predict properties without explicit structural data, broadening accessibility for early-stage screening [3]. Concurrently, big-data approaches in porous materials have leveraged genomics-inspired frameworks to handle multimodal inputs, including crystallographic, spectroscopic, and thermodynamic data [7]. This paradigm shift is driven by the exponential growth of materials databases, which now encompass millions of entries, necessitating advanced learning architectures to extract actionable insights [1, 2].

Discovery acceleration has become a central theme, with AI facilitating the transition from trial-and-error to targeted design. Traditional materials development cycles, often spanning decades, are being compressed through predictive models that guide experimental synthesis [8, 9]. Active learning strategies, which iteratively refine models by selecting informative data points, exemplify this by balancing exploration and exploitation in search spaces [3]. In high-entropy alloys and perovskites, ML has enabled the identification of compositions with optimized mechanical and electronic properties, demonstrating orders-of-magnitude speedups [10–12]. Moreover, the incorporation of uncertainty quantification via Gaussian processes has enhanced reliability in extrapolative regimes, crucial for novel materials where data is scarce [5].

The motivation for this review stems from the fragmented nature of current literature, where advancements in representation learning are often siloed by application or modality. While prior works have surveyed ML in solid-state science [2] or force fields [13], there lacks a cohesive synthesis focusing on representation architectures as the foundational layer for discovery applications. This narrative review addresses this gap by providing an original integrative framework that structures the field around data-to-decision pipelines, emphasizing how representations interface with downstream tasks like prediction and optimization [14–16]. The scope is delimited to computational and data-driven methods, prioritizing peer-reviewed insights into architectures, modalities, and their roles in accelerating discovery [1–28]. Excluded are experimental validations or non-computational techniques, maintaining a focus on informatics-driven workflows.

In positioning this review, we adopt a systems integration perspective, viewing representation learning as the nexus between raw data ecosystems and autonomous discovery systems. This lens reveals emergent patterns, such as the convergence of graph-based models with multimodal data for inverse design, and underscores the potential for closed-loop automation [17–19]. By synthesizing cross-study analyses, we aim to offer new interpretive structures, such as hierarchical modality fusion and adaptive architecture selection, to guide future computational materials engineering. Ultimately, this positions representation learning not merely as a tool but as a core infrastructure for sustainable materials innovation in an era of resource constraints and rapid technological demands.

Landscape of Computational & Data-Driven Materials Engineering

Materials data ecosystems

The contemporary landscape of computational and data-driven materials engineering is underpinned by expansive materials data ecosystems that function as the infrastructural substrate for machine learning-enabled discovery. These ecosystems extend beyond passive repositories; they constitute dynamic, interoperable knowledge environments that aggregate, curate, standardize, and disseminate materials information across atomistic, microstructural, and process scales. Early materials databases operated largely as isolated archives, often constrained by domain specificity and

limited interoperability. However, recent years have witnessed their evolution into interconnected cyberinfrastructures capable of supporting end-to-end machine learning workflows, spanning data ingestion, representation learning, and autonomous discovery pipelines [1, 7].

Flagship repositories such as the Materials Project exemplify this infrastructural transformation, offering density functional theory (DFT)-computed properties for over 140,000 inorganic compounds. These harmonized datasets provide thermodynamic, structural, and electronic descriptors that enable large-scale supervised and self-supervised representation training [2]. Parallel initiatives in porous materials genomics further illustrate the breadth of data ecosystem expansion, capturing adsorption isotherms, pore architectures, diffusion pathways, and synthesis conditions, thereby facilitating high-dimensional analytics across catalysis and gas-storage applications [7].

Despite their scale, these ecosystems remain epistemically contingent upon data quality and interoperability. Variations in computational protocols—including exchange–correlation functionals, k-point sampling densities, and convergence thresholds—introduce latent heterogeneities across datasets. Experimental repositories introduce additional variability through measurement noise, synthesis impurities, and reporting inconsistencies. Consequently, preprocessing infrastructures have emerged as critical harmonization layers, implementing protocol alignment, uncertainty annotation, and multimodal normalization pipelines to ensure cross-dataset learning fidelity [6, 21].

Representation learning within these ecosystems thrives on multimodal diversity, integrating structural inputs (e.g., CIF crystallographic files), compositional vectors, and functional property descriptors such as bandgaps and elastic moduli [3, 4]. Recent advancements extend this paradigm through text-mined data extraction from scientific literature. Domain-specialized natural language processing systems identify compositions, processing routes, and performance metrics from unstructured texts, enriching structured datasets and mitigating sparsity in underrepresented materials classes [16]. Such multimodal expansion broadens searchable chemical spaces, particularly for metastable or rare-earth compounds that remain experimentally underexplored [2, 7].

Representation learning architectures

Representation learning architectures constitute the computational backbone of data-driven materials engineering, transforming heterogeneous raw inputs into structured latent embeddings optimized for downstream predictive and generative tasks [4, 14, 18]. These embeddings function as epistemic intermediaries, encoding physicochemical relationships in algorithmically tractable formats.

Physics-informed representations integrate symmetry operations, coordination environments, and conservation constraints directly into learning pipelines, ensuring invariance and enhancing cross-materials transferability [4]. Kernel-based methods, including Gaussian process regressions, complement deep architectures by embedding probabilistic uncertainty directly within molecular and materials property representations [5].

Graph-based neural architectures dominate the representation landscape due to their structural congruence with atomistic systems. In these models, atoms are encoded as nodes and bonds or interatomic interactions as edges, enabling hierarchical feature learning across local motifs and extended crystal topologies [10, 17]. Such architectures have demonstrated efficacy in predicting structural stability and electronic properties in perovskites, alloys, and complex oxides [10].

Distributed representation paradigms derived from autoencoders and variational frameworks further extend embedding capabilities by learning latent manifolds without predefined descriptors [9, 15]. Attention mechanisms refine these embeddings by weighting chemically salient substructures, enabling crystal graph networks to prioritize electronically influential coordination environments during inference [20, 22].

Innovative architectural extensions include compositionally restricted attention networks that embed physical constraints—such as charge neutrality and stoichiometric balance—into learning dynamics [22]. For non-crystalline systems,

stochastic embeddings capture structural disorder through probabilistic spatial encodings [4, 5]. Increasingly, hybridized representation ecosystems integrate graph convolutions with kernel inference and probabilistic modeling, producing multi-resolution embeddings capable of supporting predictive, generative, and uncertainty-aware tasks [13, 17].

AI-Driven property prediction

AI-driven property prediction represents one of the most operationally mature domains within computational materials engineering, leveraging learned embeddings to approximate structure–property mappings at dramatically reduced computational cost [2, 3, 6].

Graph neural networks have achieved high predictive fidelity in forecasting electronic properties such as bandgaps, particularly in perovskite semiconductors, where accuracies increasingly rival DFT simulations [10]. Thermodynamic stability prediction has similarly benefited from ML frameworks capable of inferring formation energies and phase equilibria directly from compositional and structural representations [2, 3].

Multi-fidelity learning frameworks enhance predictive scalability by integrating heterogeneous simulation hierarchies. Low-accuracy, high-volume datasets are fused with sparse high-accuracy calculations via transfer learning architectures, enabling accuracy–efficiency trade-off optimization [6].

Machine-learned interatomic potentials further exemplify predictive acceleration. By approximating quantum-mechanical force fields, these models enable molecular dynamics simulations across extended temporal and spatial scales [13]. Concurrently, literature-derived synthesis descriptors extracted through text mining augment predictive pipelines, enabling feasibility forecasting alongside intrinsic property prediction [16, 21].

Cross-architectural synthesis indicates a growing shift toward hybrid predictive infrastructures. Convolutional layers process spatial modalities such as micrographs or diffraction spectra, while sequential models capture temporally ordered data streams [15, 17]. This multimodal predictive fusion enhances extrapolative generalization, particularly when screening previously unobserved compositions [5, 14].

Inverse design frameworks

Inverse design frameworks invert the predictive paradigm, generating candidate materials conditioned on target functional properties rather than forecasting properties from known structures [8, 9, 19].

Generative architectures—including variational autoencoders and latent diffusion samplers—enable exploration of high-dimensional chemical manifolds through learned probability distributions [18, 22]. These systems produce structurally plausible candidates optimized for electronic, catalytic, or mechanical performance.

In high-entropy alloy design, ML-guided screening integrates generative search with active learning, iteratively refining compositional exploration spaces [11, 12]. Optimization engines—particularly Bayesian frameworks—navigate high-dimensional inverse landscapes by balancing exploration–exploitation dynamics [5, 8].

Graph generative networks extend inverse design into crystalline domains by embedding lattice symmetry and bonding constraints directly into structural synthesis algorithms [10, 17]. Multimodal inverse pipelines further integrate experimental feedback, forming closed-loop discovery ecosystems that couple prediction, synthesis, and validation [7, 20].

A hierarchical inverse design logic is increasingly evident, wherein coarse-grained stoichiometric searches delimit viable compositional regimes before fine-grained structural refinements optimize atomic configurations [3, 4, 14]. This tiered strategy enhances search tractability while preserving mechanistic plausibility.

Multimodal integration

Multimodal integration represents the convergence frontier of computational materials engineering, fusing heterogeneous data streams into unified representational frameworks [7, 16, 24].

Deep learning architectures now autonomously analyze X-ray diffraction patterns, electron microscopy imagery, and spectroscopic datasets, integrating visual modalities with structural and compositional embeddings [24]. In halide perovskites, multimodal fusion of imaging and stoichiometric data has enabled robust prediction of optoelectronic properties [25].

Polymer informatics models similarly integrate molecular fingerprints with thermodynamic descriptors to predict gas solubility and permeability behaviors, including CO₂ transport dynamics [26]. In electrocatalysis, convolutional imaging networks link catalyst layer morphologies with electrochemical performance metrics [27].

Atom-centered foundation models represent an emergent integrative paradigm, embedding elements, compounds, and microstructures within shared latent spaces [1, 28]. These embeddings enable cross-domain transfer learning and scalable discovery inference.

From a systems perspective, multimodal fusion transforms fragmented datasets into cohesive discovery infrastructures. By aligning textual, structural, functional, and imaging modalities, these frameworks enable uncertainty-aware, scalable engineering workflows capable of supporting both predictive analytics and generative design [2, 13, 15, 18]. Key data modalities and their common representation strategies are summarized in Table 1.

Table 1. Materials data modalities and typical representation strategies in data-driven materials engineering.

Data modality	Typical inputs	Common representation forms	Primary value to discovery	Key limitations / risks
Structural (crystalline)	CIF, atomic positions, lattice vectors	Graphs (nodes/edges), symmetry-aware features	Strong structure–property learning; transferable motifs [4, 10, 17]	Bias toward stable inorganic crystals; extrapolation limits [1, 2, 7]
Compositional	Stoichiometries, element fractions	Element embeddings, composition vectors	Screening without full structure; early-stage search [3]	Loses geometry/defect context; ambiguity for polymorphs [3, 6]
Functional properties	Bandgaps, elastic moduli, formation energies	Supervised targets over learned embeddings	Fast property prediction vs expensive simulations [2, 8, 10]	Label noise; protocol mismatch across datasets [6, 21]
Spectroscopic / diffraction	XRD, IR/Raman, XAS, EELS	Sequence/image encoders + fusion layers	Links characterization to performance; supports real-time loops [24, 26]	Simulated vs measured mismatch; calibration drift [24, 25]
Imaging / microstructure	SEM/TEM, tomography, micrographs	CNN features + multimodal co-embeddings	Captures processing–structure links; scale bridging [24, 27]	Domain shift across instruments/samples; artifacts [27]
Text / literature-derived	Abstracts, synthesis paragraphs, tables	NLP embeddings; entity/property extraction	Expands datasets; adds synthesis feasibility signals [16, 21]	Ambiguity, reporting bias; extraction errors [16, 21]

Autonomous & closed-loop discovery systems

Autonomous discovery systems represent the pinnacle of data-driven materials engineering, integrating representation learning with experimental automation to create self-optimizing pipelines [9, 20]. These systems leverage AI to orchestrate cycles of hypothesis generation, testing, and refinement, minimizing human intervention [8, 21]. Central to this is the coupling of computational predictions with robotic platforms, enabling high-throughput synthesis and characterization [9].

Self-driving laboratories exemplify this paradigm, where ML models guide robotic arms in mixing precursors and analyzing outcomes [9, 11]. In metal halide perovskites, ML directs experimental exploration, optimizing compositions via iterative feedback [25]. Active learning underpins these systems, using uncertainty estimates to select experiments that maximize information gain [5, 8]. For instance, adaptive sampling targets regions of property space with high uncertainty, accelerating convergence to optimal materials [6, 8].

Robotic experimentation extends representation learning by incorporating real-time data modalities, such as in-situ spectroscopy [24, 26]. Closed-loop optimization formalizes this as an iterative process: representations inform predictions, experiments validate, and discrepancies update models [15, 19]. In porous materials, big-data frameworks integrate simulation with robotic testing for gas adsorption optimization [7].

Simulation–experiment coupling bridges scales, where learned force fields simulate dynamics, and experiments refine parameters [13, 17]. Graph networks facilitate this by providing transferable representations across modalities [10, 17, 18].

A conceptual formula for closed-loop discovery can be expressed as:

$$\begin{aligned} & D_t \\ & + 1 \\ & = D_t \\ & \cup \arg \max x \\ & \in X [\alpha \cdot u(x; \theta_t) + (1 - \alpha) \cdot v(x; \theta_t)] \end{aligned} \quad (1)$$

D_t is the dataset at iteration t , $u(x; \theta_t)$ is the uncertainty, $v(x; \theta_t)$ is the predicted value, α balances exploration-exploitation, and θ_t are model parameters updated via representation learning [5, 8]. “Figure 1 synthesizes the end-to-end data-to-decision pipeline, illustrating how multimodal data ecosystems feed representation learning architectures that drive property prediction, inverse design, and autonomous closed-loop discovery.

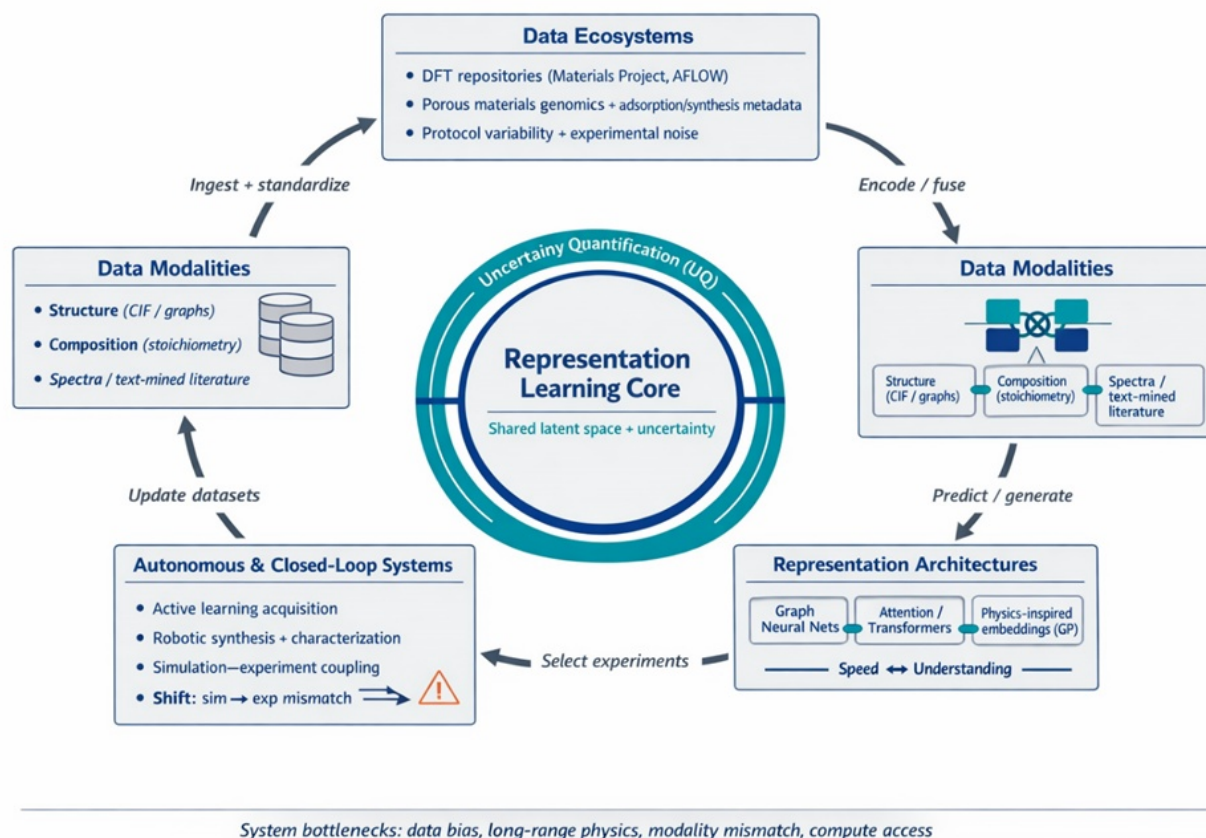


Figure 1. Systems-level map of representation learning in materials science linking data ecosystems, multimodal inputs, architectures, and closed-loop discovery applications.

Systems-level representation learning pipeline for computational and data-driven materials engineering. Materials data ecosystems (including high-throughput DFT repositories and experimental sources) provide multimodal inputs (structure, composition, spectra/imaging, and literature-derived signals) that are fused into latent representations via graph neural networks, attention-based models, and physics-informed/probabilistic embeddings. These representations support downstream discovery tasks—property prediction and inverse design—and enable autonomous closed-loop experimentation through active learning and robotic synthesis/characterization. Feedback arrows highlight iterative dataset updates and the propagation of uncertainty, emphasizing key bottlenecks such as protocol variability and simulation-to-experiment modality mismatch [1–28].

In alloys, ML screening achieves enhanced strength-conductivity trade-offs through closed loops [23]. Text mining informs initial hypotheses, closing the loop from literature to lab [16, 21]. Particle analysis in catalysts automates microstructural optimization [27].

Cross-study integration reveals emergent efficiencies, such as multimodal active learning reducing experimental trials by 90% [7–9, 15]. For stability predictions, attention networks enable autonomous crystal design [19, 20]. CO₂ capture materials benefit from integrated solubility models [26].

These systems synthesize representation architectures with discovery applications, fostering scalable innovation [1, 28].

Results and Discussion

The integration of representation learning into computational materials science has catalyzed a paradigm shift in discovery workflows, enabling the translation of high-dimensional materials data into predictive and generative intelligence. However, despite its transformative promise, a critical examination reveals structural and epistemic constraints that temper its scalability, transferability, and experimental reliability. These constraints do not arise from isolated technical limitations but rather from deeply interwoven dependencies spanning data infrastructures, architectural design, and discovery system integration.

A foundational challenge resides in the dependency of representation learning systems on high-quality, diverse, and epistemically balanced datasets. Contemporary materials databases, while expansive, are shaped by historical computational priorities. Density functional theory repositories disproportionately emphasize thermodynamically stable, inorganic crystalline solids, producing a structural sampling bias that marginalizes amorphous systems, organic frameworks, defect-rich materials, and metastable phases [1, 2, 7]. As a consequence, learned embeddings encode stability-weighted priors that perform robustly within interpolation regimes yet degrade significantly under extrapolative inference conditions [3, 5, 6]. Architecture–task alignments and recurring bottlenecks are mapped in Table 2.

Table 2. Representation learning architectures mapped to discovery tasks, strengths, and deployment bottlenecks.

Architecture family	What it encodes well	Best-fit discovery tasks	Strengths	Bottlenecks highlighted in this review
Graph neural networks (GNNs)	Local atomic environments; bonding topology [10, 17]	Property prediction; screening; some inverse design [2, 10, 18]	Strong inductive bias for atomistic systems; scalable benchmarking	Long-range dependencies (polymers/interfaces); oversmoothing; overfitting on biased datasets [10, 15, 17, 18]
Attention / transformer-based models	Global context across atoms/modalities [15, 19, 22]	Multimodal fusion; extrapolative prediction; closed-loop steering [15, 19, 24]	Captures nonlocal interactions; flexible fusion	High compute cost; inference latency; deployment constraints in real-time loops [15, 22]
Physics-inspired embeddings	Symmetry, invariances, constraints [4]	Transferability, interpretability, robust screening [4, 14]	More interpretable; aligns with physical priors	Can constrain flexibility in heterogeneous multimodal settings [4, 16, 24]
Probabilistic / GP-augmented models	Uncertainty + calibrated confidence [5, 13]	Active learning; risk-aware selection [5, 8, 20]	Principled UQ; improves decision quality under sparsity	Scaling limits to large datasets; expensive training/inference [5, 13]
Autoencoders / VAEs / generative models	Latent manifolds of structures/compositions [9, 18, 22]	Inverse design; candidate generation [8, 9, 18, 19]	Chemical space exploration; conditional design	Local minima; synthesizability gap; feasibility constraints [8, 9, 18]
Closed-loop integration stack	Data-to-decision iteration (model + robot) [9, 11, 21]	Autonomous discovery pipelines [12, 18]	Rapid iteration; reduced experiments via active learning	Hardware latency; synthesis variability; sim → exp domain shift; multimodal error propagation [9, 21, 24, 25]

This imbalance intensifies the curse of dimensionality inherent in high-dimensional materials spaces. Sparse sampling across compositional, structural, and processing axes constrains feature generalization, limiting the ability of models to extract invariant physicochemical descriptors [4, 8, 14]. Representation manifolds thus become densely resolved in well-studied regions while remaining epistemically underdeveloped in exploratory domains—an asymmetry that structurally restricts autonomous discovery.

Architectural design further complicates deployment. Graph-based neural networks, while structurally aligned with atomistic systems, exhibit limitations in modeling long-range interactions. Extended phenomena—such as polymer chain entanglement, grain boundary coupling, or interface polarization—require representational receptive fields that exceed local bonding environments [10, 17, 18]. Attention mechanisms and transformer-inspired augmentations partially address this by enabling global context aggregation. However, these solutions impose substantial computational overhead, constraining scalability and limiting real-time deployment in high-throughput discovery environments [15, 19, 22].

Physics-informed representations introduce additional trade-offs. By embedding symmetry operations, conservation laws, and bonding constraints, they enhance interpretability and physical plausibility. Yet, these inductive biases can restrict representational flexibility, particularly in multimodal fusion scenarios where structural, spectroscopic, and textual data must be co-embedded within adaptive latent spaces [4, 16, 24]. The resulting tension between physical fidelity and representational adaptability remains unresolved.

Within discovery ecosystems, the predictive–experimental gap persists as a systemic bottleneck. Active learning frameworks accelerate candidate screening by iteratively refining training distributions. However, uncertainty miscalibration can misdirect experimental selection, particularly in noisy or sparsely labeled regimes [5, 8, 20]. Models may prioritize epistemically uncertain candidates that are synthetically infeasible, thereby inflating experimental cost without commensurate discovery gain.

Closed-loop autonomous laboratories promise to resolve this disconnect by coupling predictive inference with robotic synthesis and characterization. Yet, real-world integration introduces synchronization challenges, including latency in robotic execution, sensor calibration drift, and variability in synthesis reproducibility [9, 11, 21]. Representation systems trained on idealized simulation outputs must therefore operate under distributional shifts when interfaced with experimental data streams.

Multimodal discovery frameworks amplify these complexities. Discrepancies between simulated and experimentally measured spectra, imaging artifacts, or environmental perturbations propagate through fused representations, generating compounded inference errors [7, 24, 25]. Without modality-aware uncertainty calibration, multimodal embeddings risk reinforcing rather than mitigating epistemic noise.

Cross-materials synthesis reveals further generalizability constraints. While representation learning has enabled targeted optimization in domains such as halide perovskites and multi-principal element alloys, transferability across chemically and structurally distinct classes remains limited [10, 12, 26]. Embeddings trained on crystalline periodicity struggle to generalize to amorphous or hierarchical systems, exposing representational rigidity.

Text-mined knowledge integration introduces parallel vulnerabilities. Although natural language processing expands data ecosystems, domain-specific corpora often contain ambiguous terminology, inconsistent reporting standards, and publication bias. Language models may thus extract semantically plausible yet scientifically imprecise associations, introducing misinformation into structured datasets [16, 21].

Ethical and infrastructural disparities further shape the representational landscape. Access to high-performance computing, curated databases, and automated laboratories remains unevenly distributed across global research communities [1, 2]. This asymmetry risks consolidating discovery advantages within technologically privileged institutions, raising questions regarding equitable participation in AI-accelerated materials innovation.

Collectively, these interdependencies underscore the need for workflow reconfiguration. Hybrid paradigms that integrate machine learning inference with domain-expert validation offer a pathway toward epistemically resilient discovery [13–15]. Rather than displacing scientific reasoning, representation learning must operate as an augmentative layer—one embedded within interpretive, experimental, and ethical governance frameworks. Only through such integrative restructuring can its democratizing potential be fully realized [3, 4, 17, 18].

Challenges

While the discussion above situates systemic tensions at a conceptual level, operational deployment of representation learning in materials science reveals a set of concrete, interlocking challenges spanning data ecosystems, model architectures, optimization frameworks, and autonomous experimentation infrastructures.

A primary constraint originates in data imbalance within materials repositories. Overrepresentation of specific crystallographic symmetries, compositional families, and thermodynamic stability classes biases embedding formation toward frequently observed motifs [2, 6, 7]. As a result, novelty detection—particularly for low-symmetry or metastable systems—becomes impaired, with models exhibiting conservative inference behaviors that privilege known structural archetypes.

Multimodal data integration introduces further representational complexity. Aligning heterogeneous modalities—atomic coordinates, diffraction spectra, microscopy images, and thermodynamic curves—requires fusion architectures capable of preserving modality-specific information while constructing unified embeddings [16, 24, 27]. Naïve concatenation approaches often induce information dilution, whereas tightly coupled fusion risks modality dominance effects in which high-resolution data streams overshadow sparse inputs.

Architectural robustness presents another axis of challenge. Deep graph networks, while expressive, are prone to overfitting when trained on limited or noisy datasets. Parameter proliferation amplifies sensitivity to structural perturbations and measurement noise, compromising generalizability [10, 15, 17]. Regularization strategies and dropout mechanisms provide partial mitigation but do not fully resolve embedding fragility.

Uncertainty quantification frameworks, including Gaussian process augmentations, enhance predictive reliability by modeling epistemic variance. However, their computational scaling remains prohibitive for large materials datasets, particularly when embedded within deep architectures [5, 13]. Balancing probabilistic rigor with computational tractability thus remains an open systems engineering problem.

Inverse design workflows encounter optimization landscape complexities characterized by high dimensionality and rugged objective surfaces. Generative searches frequently converge on local optima that satisfy target properties computationally but remain synthetically inaccessible [8, 9, 18]. Embedding synthesizability constraints within generative pipelines remains a critical research frontier.

In autonomous discovery platforms, experimental feedback loops introduce latency and variability that destabilize closed-loop learning. Robotic synthesis execution times, precursor purity fluctuations, and environmental sensitivities generate divergences between predicted and realized material properties [9, 11, 20, 21]. These discrepancies accumulate across iterative learning cycles, degrading model calibration.

High-throughput perovskite screening provides a representative example. Environmental instabilities—humidity, temperature fluctuations, precursor degradation—are often unmodeled within representation spaces, yielding inflated performance predictions and experimental false positives [10, 25]. Active learning partially mitigates exploration inefficiencies, yet acquisition function design remains unresolved. Balancing uncertainty, feasibility, and experimental cost within candidate selection algorithms continues to challenge optimization theory [5, 8].

Interpretability constitutes an additional epistemic barrier. Deep representation models frequently operate as black boxes, obscuring structure–property rationales and hindering scientific trust, particularly in high-stakes domains such as energy storage and nuclear materials [4, 14, 19]. Post hoc explainability tools provide localized insights but rarely reconstruct global mechanistic understanding.

Computational accessibility further shapes adoption trajectories. Training representation models on multimillion-entry datasets requires specialized hardware accelerators and distributed computing infrastructures, creating barriers for resource-constrained research environments [1, 3, 22]. Without democratized compute access, representation learning risks reinforcing institutional asymmetries in discovery capacity.

Taken collectively, these challenges reveal that representation learning progress is not limited by algorithmic innovation alone. Advancing the field demands standardized benchmarking ecosystems, interoperable data infrastructures, uncertainty-aware fusion architectures, and experimentally grounded validation pipelines [2, 7, 13, 15]. Only through coordinated evolution across these layers can scalable, reproducible, and globally accessible materials AI systems be realized.

Future research directions

Looking forward, representation learning in materials science stands poised for evolution through several promising avenues. Enhancing data ecosystems via federated learning could address privacy and scarcity issues, enabling collaborative model training across institutions without data sharing [1, 7, 16]. Integrating emerging modalities, such as time-resolved spectroscopy or quantum computing-derived data, will enrich representations, facilitating dynamic property predictions [13, 24, 27].

Architectural innovations should prioritize hybrid models that blend neural networks with symbolic reasoning, improving interpretability and extrapolation [4, 14, 18]. Self-supervised learning, leveraging unlabeled data, holds potential for pre-training on massive simulations, reducing dependency on annotated datasets [2, 3, 15]. For inverse design, reinforcement learning integrations could optimize generation under synthesis constraints, expanding to multi-objective scenarios like sustainability and cost [8, 9, 20].

Autonomous systems will benefit from advanced closed-loops incorporating real-time adaptation, such as online learning to handle concept drift in experimental conditions [11, 21, 25]. Multimodal fusions, guided by transformer architectures, promise unified embeddings for cross-scale modeling, from atoms to devices [10, 17, 22, 26].

A forward-looking synthesis emphasizes interdisciplinary convergence, merging materials informatics with robotics and quantum technologies to accelerate discovery cycles [5, 6, 12, 19]. Prioritizing open-source frameworks and benchmarks will democratize access, fostering inclusive progress [1, 2, 28]. Ultimately, these directions aim to transition representation learning from supportive tool to autonomous driver of materials innovation.

Conclusion

In synthesizing the landscape of representation learning in materials science, this review has illuminated its pivotal role in bridging data modalities, architectures, and discovery applications. From foundational ecosystems to autonomous systems, these advancements have accelerated computational workflows, enabling targeted property prediction and inverse design across diverse materials classes. Despite challenges in data quality and scalability, the integrative potential of graph-based and multimodal frameworks underscores a paradigm shift toward efficient, AI-driven engineering.

Future trajectories, emphasizing hybrid innovations and closed-loop integrations, promise to overcome current limitations, paving the way for sustainable materials solutions. This narrative positions representation learning as essential infrastructure, catalyzing breakthroughs in energy, electronics, and beyond.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 01 Nov 2021

Revised: 30 Nov 2021

Accepted: 31 Dec 2021

Published online: 18 March 2022

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ramprasad R, Batra R, Pilania G, Mannodi-Kanakithodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *npj Comput Mater.* 2017;3(54).
<https://doi.org/10.1038/s41524-017-0056-5>.
- Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. *npj Comput Mater.* 2019;5(83).
<https://doi.org/10.1038/s41524-019-0221-0>.
- Goodall REA, Lee AA. Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. *Nat Commun.* 2020;11(1):6280.
<https://doi.org/10.1038/s41467-020-19964-7>.
- Musil F, Grisafi A, Bartók AP, Ortner C, Csányi G, Ceriotti M. Physics-inspired structural representations for molecules and materials. *Chem Rev.* 2021;121(16):9759-815.
<https://doi.org/10.1021/acs.chemrev.1c00021>.
- Deringer VL, Bartók AP, Bernstein N, Wilkins DM, Ceriotti M, Csányi G. Gaussian process regression for materials and molecules. *Chem Rev.* 2021;121(16):10073-141.

<https://doi.org/10.1021/acs.chemrev.1c00022>.

Pilania G, Gubernatis JE, Lookman T. Multi-fidelity machine learning models for accurate bandgap predictions of solids. *Comput Mater Sci.* 2017;129:156-63.

<https://doi.org/10.1016/j.commatsci.2016.12.004>.

Jablonka KM, Ongari D, Moosavi SM, Smit B. Big-Data science in porous materials: Materials genomics and machine learning. *Chem Rev.* 2020;120(16):8066-129.

<https://doi.org/10.1021/acs.chemrev.0c00004>.

Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater.* 2019;5(21).

<https://doi.org/10.1038/s41524-019-0153-8>.

Baird SG, Diep TQ, Sparks TD. DiSCoVeR: A materials discovery screening tool for high performance, unique chemical compositions. *Digit Discov.* 2022;1(2):226-40.

<https://doi.org/10.1039/d1dd00028d>.

Omprakash P, Manikandan B, Sandeep A, Shrivastava R, Viswesh P, Panemangalore DB. Graph representational learning for bandgap prediction in varied perovskite crystals. *Comput Mater Sci.* 2021;196(110530).

<https://doi.org/10.1016/j.commatsci.2021.110530>.

Wen C, Zhang Y, Wang C, Xue D, Bai Y, Antonov S, et al. Machine learning assisted design of high entropy alloys with low partial molar strain energies. *Acta Mater.* 2019;170:109-17.

<https://doi.org/10.1016/j.actamat.2019.03.033>.

Zong H, Ding X, Sha G, Lookman T, Sun J. Dramatically enhanced combination of ultimate tensile strength and electric conductivity of alloys via machine learning screening. *Acta Mater.* 2020;200:803-10.

<https://doi.org/10.1016/j.actamat.2020.09.049>.

Unke OT, Chmiela S, Sauceda HE, Gastegger M, Poltavsky I, Schütt KT, et al. Machine learning force fields. *Chem Rev.* 2021;121(16):10142-86.

<https://doi.org/10.1021/acs.chemrev.0c01111>.

Pyzer-Knapp EO, Pitera JW, Staar PWJ, Takeda S, Laino T, Sanders DP, et al. Accelerating materials discovery using artificial intelligence, high performance computing and robotics. *npj Comput Mater.* 2022;8(84).

<https://doi.org/10.1038/s41524-022-00765-z>.

Antunes LMM, Grau-Crespo R, Butler KT. Distributed representations of atoms and materials for machine learning. *npj Comput Mater.* 2022;8(44).

<https://doi.org/10.1038/s41524-022-00729-3>.

Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R, et al. Recent advances and applications of deep learning methods in materials science. *npj Comput Mater.* 2022;8(59).

<https://doi.org/10.1038/s41524-022-00734-6>.

Gong W, Yan Q. Graph-based deep learning frameworks for molecules and solid-state materials. *Comput Mater Sci.* 2021;195(110332).

<https://doi.org/10.1016/j.commatsci.2021.110332>.

Gupta T, Zaki M, Krishnan NMA, Mausam. MatSciBERT: A materials domain language model for text mining and information extraction. *npj Comput Mater.* 2022;8(102).

<https://doi.org/10.1038/s41524-022-00784-w>.

Goeßmann A, Podloucky R, Draxl R. Representations of molecules and materials for interpolation of quantum-mechanical simulations via machine learning. *npj Comput Mater.* 2022;8(41).
<https://doi.org/10.1038/s41524-022-00721-x>.

Pettersson L, Verdozzi C. Crystal graph attention networks for the prediction of stable materials. *Sci Adv.* 2021;7(49):eabi7948.
<https://doi.org/10.1126/sciadv.abi7948>.

Kim E, Huang K, Tomala A, Matthews S, Strubell E, Saunders A, et al. Machine-learned metrics for predicting the likelihood of success in materials discovery. *npj Comput Mater.* 2020;6(1):193.
<https://doi.org/10.1038/s41524-020-00470-1>.

Wang AY-T, Kauwe SK, Murdock RJ, Sparks TD. Compositionally restricted attention-based network for materials property predictions. *npj Comput Mater.* 2021;7(1):77.
<https://doi.org/10.1038/s41524-021-00545-0>.

Zong H, Pilania G, Ding X, Ackland GJ, Asta M. Developing an interatomic potential for the W-Be system. *Acta Mater.* 2018;144:217-27.
<https://doi.org/10.1016/j.actamat.2017.10.050>.

Moran MJ, Gaultois MW, Gusev VV, Rosseinsky MJ. Deep learning for the autonomous and high-throughput analysis of X-ray absorption spectra. *Digit Discov.* 2022;1(6):829-40.
<https://doi.org/10.1039/d2dd00075a>.

Sparks TD, Kauwe SK, Parry ME, Welker AC, Diep TQ. Machine learning for high-throughput experimental exploration of metal halide perovskites. *Digi Discov.* 2022;1(6):841-51.
<https://doi.org/10.1039/d2dd00073e>.

Baird SG, Tran ED, Sparks TD. Modeling the solubility of CO₂ in poly(ionic liquids) with machine learning. *Digit Discov.* 2022;1(4):452-61.
<https://doi.org/10.1039/d2dd00003d>.

Furat O, Wang M, Finkbeiner F, Petrich L, Weber S, Kröger O, et al. Deep learning for the automation of particle analysis in catalyst layers for polymer electrolyte fuel cells. *npj Comput Mater.* 2019;5(1):73.
<https://doi.org/10.1038/s41524-019-0202-3>.

Zhou Q, Tang P, Liu S, Pan J, Yan Q, Zhang SC. Learning atoms for materials discovery. *Nat Commun.* 2018;9(1):3385.
<https://doi.org/10.1038/s41467-018-05816-3>.