ORIGINAL RESEARCH **Open access**

Compositional Generalization as a Distinct Failure Mode in Materials AI

Daniel Fischer[1*], Laura Meier[1], Thomas Braun[2]

Abstract

The integration of artificial intelligence into materials science has highlighted challenges in model performance, particularly in domains that require extrapolation beyond the training data distribution. This manuscript explores compositional generalization as a unique failure mode in materials AI, in which systems struggle to interpret novel combinations of atomic or molecular elements despite familiarity with individual components. Through a synthesis of recent literature, the analysis delineates how this failure manifests in predictive tasks, such as property estimation in alloys or polymers, revealing underlying tensions between data-driven learning and structural comprehension. Conceptual interpretations highlight the interplay between representational invariance and contextual dependencies, underscoring epistemic gaps in current architectures. The proposed framework interprets these dynamics through lenses of modular interaction and systemic feedback, emphasizing trade-offs in scalability and robustness. By examining the ethical ramifications of deployment in high-stakes applications, the discussion integrates insights into steering mechanisms that could mitigate such limitations without empirical validation. Ultimately, this conceptual inquiry fosters a deeper understanding of AI's role in advancing materials discovery and advocates for interpretive strategies that prioritize holistic integration over isolated optimizations.

Keywords Materials AI, Conceptual framework, Compositional generalization, Failure modes, Epistemic challenges, Systemic interactions

*Correspondence:
Daniel Fischer
daniel.fischer@outlook.com
[1] Department of Materials Science and Intelligent Systems, Faculty of Engineering, University of Freiburg, Freiburg, Germany
[2] Department of AI-Driven Materials Engineering, Faculty of Engineering, Karlsruhe Institute of Technology, Karlsruhe, Germany

Introduction

The advent of artificial intelligence (AI) in materials science has reshaped the landscape of discovery and design, enabling accelerated exploration of vast chemical spaces that traditional methods could scarcely navigate. At its core, materials AI leverages machine learning to infer patterns from datasets of atomic structures, properties, and behaviors, thereby enabling predictions that inform synthesis and application. However, amid these advancements, persistent challenges emerge, particularly in the realm of generalization— the capacity of models to extend learned knowledge to unseen scenarios. Among these, compositional generalization stands out as a particularly insidious failure mode, in which AI systems falter at recombining familiar elements into novel configurations, a process integral to materials innovation [1, 2].

In materials contexts, compositional generalization refers to a model's ability to handle new mixtures or arrangements of known constituents, such as predicting the mechanical properties of an alloy composed of elements that are individually well-represented in the training data but combined in unprecedented ratios. This failure mode differs from other generalization issues, such as out-of-distribution shifts in environmental conditions or scale mismatches, by specifically targeting the combinatorial essence of material composition [3, 4]. For instance, while a model might accurately predict

properties for binary systems, it often underperforms when extrapolating to ternary or higher-order systems, revealing a disconnect between component-wise learning and holistic synthesis [5, 6].

The roots of this challenge lie in the foundational assumptions underlying the AI architectures used in materials science. Graph neural networks (GNNs), convolutional networks, and transformer-based models, which are commonly applied to atomic-scale representations, prioritize local interactions but may overlook emergent behaviors arising from global composition [7, 8]. This architectural bias amplifies epistemic uncertainty, as the model's internal representations fail to capture the invariant principles governing material behavior across compositional variations [9, 10]. Consequently, deployment in real-world scenarios, such as the design of high-entropy alloys or functional polymers, risks propagating errors that could undermine reliability in sectors such as energy storage or aerospace [11, 12].

Moreover, the data ecosystems underpinning materials AI exacerbate this failure. Databases like the Materials Project or ICSD provide extensive but often biased coverage, favoring common compositions while undersampling rare or extreme combinations [1, 13]. This distributional skew fosters models that excel in interpolation but struggle with extrapolation, mirroring broader AI pitfalls observed in natural language processing or computer vision, yet uniquely manifested in the physico-chemical domain [2, 3]. Interpretive analyses suggest that such limitations stem from an overreliance on statistical correlations rather than causal or physical priors, leading to brittle performance when compositional novelty disrupts expected patterns [4, 5].

Ethically, the implications of compositional generalization failures extend beyond technical confines, influencing decision-making in sustainable materials development. Inaccurate predictions could misguide resource allocation, perpetuating inefficiencies or environmental impacts in manufacturing processes [6, 7]. Systems-level insights reveal feedback loops wherein model outputs inform experimental designs, potentially reinforcing data biases if failures are not addressed conceptually [8, 9]. Thus, understanding this failure mode requires an integrative approach that blends insights from computational materials science with AI theory to uncover steering logics that enhance robustness [10, 11].

Historically, materials science has evolved from empirical trial-and-error to physics-based simulations, with AI representing the latest paradigm shift. Yet, as adoption accelerates, the need for conceptual scrutiny intensifies, particularly regarding failure modes that could stall progress [12, 13]. This manuscript posits compositional generalization not merely as a technical hurdle but as a lens through which to examine the interplay of representation, learning, and application in materials AI [1, 2]. By synthesizing recent advancements, it aims to delineate the conceptual contours of this issue, paving the way for interpretive frameworks that prioritize dynamic interactions over static predictions [3, 4]. To clarify why compositional generalization constitutes a distinct failure mode rather than a variant of conventional generalization challenges, **Table 1** contrasts it with related failure categories commonly discussed in materials AI.

Table 1. Distinguishing compositional generalization from related generalization failures in materials AI

Failure mode	Defining characteristic	Primary source of breakdown	Why is it epistemically distinct	Representative materials contexts
Compositional generalization	Failure to infer properties of novel combinations of known components	Inadequate representation of interaction effects and emergent structure	Exposes limits of modular abstraction under combinatorial expansion	High-entropy alloys, copolymers, multicomponent oxides
Distributional shift	Input data differs statistically from the training distribution	Dataset bias or environmental mismatch	Primarily statistical rather than combinatorial	Processing condition changes, temperature regimes
Scale generalization	The model fails when extrapolating across	Resolution mismatch between training and	Linked to multiscale modeling assumptions	Micro- to macro-property transfer

	length or time scales	deployment		
Structural out-of-distribution	Novel crystal structures or topologies	Absence of structural motifs in training data	Relates to representational coverage, not recombination	New lattice types, defects
Noise-induced degradation	Performance loss due to measurement or label noise	Data quality limitations	Does not probe representational synthesis	Experimental uncertainty

In the ensuing sections, a theoretical background is synthesized from the literature on AI generalization in materials, highlighting emergent themes in failure analysis [5, 6]. This foundation informs a proposed conceptual framework that interprets compositional challenges through modular and feedback-oriented perspectives, offering analytical implications for future integrations. Through this scholarly lens, the discourse seeks to enrich the epistemic foundation of materials AI, fostering a more nuanced appreciation of its limitations and potentials.

Theoretical Background and Literature Synthesis

Foundations of generalization in materials AI

Generalization in AI, broadly construed, refers to the extension of learned patterns to novel instances. This principle assumes particular significance in materials science due to the effectively unbounded combinatorial space of possible compositions, structures, and processing conditions. Within this domain, models must navigate a persistent tension between data fidelity and extrapolative capacity, because compositional elements—atoms, molecules, or phases—interact through multiscale couplings that resist simple additive assumptions and often yield emergent behaviors not linearly recoverable from local patterns [1, 3]. Recent literature emphasizes that many machine learning paradigms remain implicitly optimized for interpolation within familiar distributions, producing high benchmark performance while exhibiting fragile behavior under compositional novelty; conceptually, this has redirected interpretive attention toward representational adequacy rather than raw predictive accuracy [5, 7].

Conceptual interpretations further suggest that generalization failures frequently arise from mismatches between training distributions and real-world diversity, particularly in the high-dimensional spaces characteristic of materials properties. In such settings, sparse sampling can cause models to learn “local correctness” without acquiring stable global regularities, leading to epistemic gaps that appear only when the model is asked to reason beyond its learned neighborhood [9, 11]. In alloy design, for example, models trained on limited or clustered datasets may capture local atomic environments effectively but fail to integrate global compositional effects or long-range constraints, resulting in confident yet unreliable predictions that signal unacknowledged uncertainty rather than robust understanding [2, 13]. This dynamic foregrounds the interpretive trade-offs of model complexity: deeper architectures can enhance feature extraction, but they may also amplify overfitting to compositional regularities and database-specific biases, making “generalization” appear stronger than it is by reinforcing dominant patterns [4, 6].

Systems-level insights reveal how feedback structures within AI pipelines can perpetuate, or even harden, these weaknesses. Iterative model refinement—especially when guided by simulation outputs or model-directed data acquisition—risks entrenching compositional blind spots when initial datasets lack diversity, thereby producing self-confirming loops in which the model repeatedly “learns” from its own induced trajectory [8, 10]. Ethical reasoning in this context stresses the importance of transparency in acknowledging such limitations, ensuring that AI-assisted discovery does not inadvertently skew research priorities toward overrepresented material classes or reinforce institutional incentives that reward narrow performance metrics over epistemic robustness [1, 12].

Emergent challenges in compositional handling

As AI applications in materials science expand, compositional handling becomes a focal point of challenge, where recombination of known elements into unseen configurations tests the boundaries of learned invariances and exposes what a model treats as “transferable structure” versus “memorized correlation” [2, 4]. Recent syntheses indicate that graph-based representations, prevalent in atomic-scale modeling, often privilege connectivity and local neighborhoods over combinatorial flexibility, producing interpretations that emphasize contextual dependence: what generalizes in one compositional regime may collapse in another when global constraints shift [6, 8]. In polymeric systems, sequence variations can yield drastically different properties. Yet, models may struggle to extrapolate beyond the motifs they were trained on, revealing interaction dynamics that favor motif memorization and narrow pattern extension rather than principled compositional reasoning [10, 12].

Analytical implications suggest that this failure mode often stems from an imbalance in representational emphasis. Elemental descriptors may encode individual attributes (e.g., electronegativity, radius, valence tendencies), while neglecting synergistic effects arising only through relational, cooperative, or competitive interactions among constituents [1, 3]. Steering logics in the literature frequently advocate hybrid approaches that incorporate physical constraints or mechanistic priors, but, conceptually, these proposals can be read less as prescriptive “fixes” and more as attempts to rebalance scalability with robustness—shifting from performance-first optimization toward interpretive stability under compositional shifts [5, 7]. Epistemic reasoning extends this perspective by positioning uncertainty quantification as more than a statistical add-on: compositional uncertainty can be interpreted as a diagnostic signal of deeper knowledge deficits in AI architectures and training regimes, rather than mere noise to be suppressed [9, 11]. **Table 2** synthesizes how compositional generalization failures manifest across major materials domains, highlighting recurring epistemic patterns despite domain-specific differences.

Table 2. Manifestations of compositional generalization failure across material domains

Materials domain	Typical predictive task	Observed failure pattern	Underlying epistemic tension	Interpretive consequence
Alloys	Mechanical or thermodynamic property prediction	Accurate binary predictions, poor higher-order extrapolation	Local accuracy vs. global compositional coherence	Overconfidence in multicomponent design
Polymers	Property–sequence relationships	Memorization of motifs, failure on new sequences	Token-level learning vs. emergent behavior	Misrepresentation of structure–property links
Solid-state electrolytes	Ionic conductivity estimation	Inconsistent extrapolation under dopant variation	Descriptor sufficiency vs. pathway complexity	Unreliable screening decisions
Functional oxides	Phase stability prediction	Sensitivity to minor compositional shifts	Abstraction vs. contextual dependency	Fragile phase diagrams
High-entropy materials	Performance optimization	Collapse under combinatorial explosion	Scalability vs. interpretive robustness	Inflated expectations of AI coverage

Moreover, in complex materials such as solid-state electrolytes, compositional variations influence transport properties through intricate pathways that are shaped by defects, microstructure, and coupled ionic–electronic effects; conceptual work has therefore explored modular decompositions as interpretive strategies for handling partial generalization without assuming holistic mastery [2, 13]. Taken together, these insights foreground compositional generalization failures as manifestations of broader tensions between data-driven empiricism and theoretical grounding—tensions that remain central to how materials AI is interpreted, trusted, and deployed [4, 6].

Integration of AI architectures and materials paradigms

The convergence of modern AI architectures with materials paradigms has spurred interpretive analyses of how neural systems encode compositional information and how those encodings shape scientific reasoning. Transformer models adapted for sequential or tokenized composition data offer enhanced attention mechanisms that can, in principle, capture context-dependent interactions; however, they may also introduce vulnerabilities to positional or encoding biases, where representational choices subtly shape what the model treats as salient compositional structure [1, 5]. Literature syntheses further describe systemic feedback in which model outputs influence dataset augmentation and exploration strategies; this can mitigate generalization lapses by diversifying coverage, yet it can also exacerbate them if augmentation reinforces dominant patterns learned from already-biased datasets [7, 9].

Conceptual interpretations emphasize ethical dimensions of deploying such systems in critical applications, where compositional misjudgments could affect safety, reliability, or downstream technological decisions in structural, energy, or functional materials contexts [11, 13]. Trade-offs between computational efficiency and comprehensive coverage reinforce the need for integrative strategies that frame AI as one component of a broader epistemic ecosystem, co-determined by human expertise, institutional incentives, and the interpretive norms of the field [2, 4]. In quantum-assisted frameworks, the interplay between classical learning and quantum-derived descriptors offers an additional lens for understanding extrapolative limits, interpreting failures not only as deficits but as prompts to refine representational schemas and re-evaluate assumptions about what “generalizes” across compositional regimes [6, 8]. Systems-level insights thus advocate for dynamic adaptations that account for evolving material challenges, fostering resilience against compositional shifts while retaining epistemic humility about what models can meaningfully claim [10, 12].

Cross-disciplinary insights and epistemic considerations

Insights from broader AI research domains provide a productive lens for interpreting compositional generalization failures in materials science. In particular, analogies to linguistic compositionality illuminate how semantic meaning emerges from the structured recombination of tokens, paralleling how material properties arise from the interaction of elemental constituents rather than from their isolated attributes alone [2, 6]. In natural language processing, models often succeed at syntactic recombination while failing at semantic generalization under distributional shift. This pattern closely mirrors how materials AI models may recombine known elements while misrepresenting emergent behaviors in novel compositions. This cross-pollination enriches conceptual frameworks in materials AI by reframing generalization failures not as anomalies but as expected consequences of representational assumptions inherited from adjacent AI paradigms.

Drawing further on transfer learning and domain adaptation literatures, recent syntheses interpret compositional failures as symptoms of insufficient invariance rather than inadequate data volume alone [8, 10]. From this perspective, models generalize not because they have seen “enough” examples, but because they have learned representations that preserve meaningful structure under transformation. When such invariances are poorly aligned with the generative processes of materials systems, extrapolation degrades—even when performance metrics remain high within familiar regimes. These interpretations emphasize that compositional generalization is fundamentally an epistemic challenge: it reflects how models encode assumptions about what aspects of structure, interaction, or context are transferable across regimes.

Epistemic reasoning deepens this view by foregrounding the interpretive value of failure itself. Rather than treating generalization breakdowns as errors to be suppressed, recent conceptual work frames them as diagnostic signals that expose hidden assumptions in model training, dataset construction, and representational abstraction [1, 12]. In this sense, compositional generalization failures act as epistemic probes, revealing where AI architectures implicitly rely on correlations that lack causal or mechanistic grounding. Ethical trade-offs emerge at this juncture, as the pressure to accelerate discovery can conflict with the responsibility to ensure reliability and interpretive transparency. Balancing innovation speed with epistemic caution becomes especially salient in materials contexts tied to sustainability, safety, or long-term infrastructure, where misplaced confidence can propagate downstream risks [3, 5].

Integrative perspectives synthesize these strands by situating compositional challenges within broader debates in AI ethics and governance. Here, steering mechanisms are interpreted not as control levers for optimization, but as reflective

practices that prioritize holistic understanding over fragmented prediction. Compositional generalization thus becomes emblematic of a larger ethical concern: whether AI systems are designed to support scientific sense-making or merely to extend narrow performance regimes. From this vantage, materials AI inherits the same ethical tensions observed across AI systems—between scalability and accountability, abstraction and meaning, efficiency and epistemic responsibility [2, 4, 6, 7, 9, 11, 13].

Proposed conceptual framework

Addressing compositional generalization as a failure mode in materials AI requires an interpretive framework that treats it as an emergent property of interacting representational and systemic components rather than an isolated modeling deficiency. The proposed framework conceptualizes compositional generalization through modular interactions coupled with feedback-driven integration layers, emphasizing how local encodings and global synthesis co-evolve within AI pipelines. At its core, the framework interprets failure as arising from tensions between representational modules—designed to encode individual components with precision—and integrative layers that synthesize novel combinations under uncertainty.

Rather than positing discrete solutions, the framework foregrounds analytical implications of how these modules interact, highlighting trade-offs between fidelity and flexibility. In alloy prediction, for example, modular representations may accurately encode elemental descriptors or local environments, yet when embedded within broader compositional contexts, feedback effects can distort estimates of emergent properties. These distortions do not simply reflect data scarcity; they expose epistemic tensions in scaling from binary or ternary systems to high-order compositional spaces, where interaction effects multiply, and abstraction choices become consequential [3, 4, 7, 8].

Within this framework, steering logics are interpreted as navigational principles that balance local optimization against global coherence. Interaction dynamics are modeled as recursive feedback loops: initial component encodings influence emergent property predictions, which in turn shape confidence, uncertainty estimates, and downstream data selection. In unseen regimes, these loops may amplify uncertainty or, conversely, mask it through overconfident abstraction [2, 3, 11, 12]. Ethical reasoning integrates directly into this structure by examining how such feedback loops influence decision-making under constraint. Overreliance on modular outputs—particularly when uncertainty signals are muted—can misdirect material prioritization, resource allocation, or experimental focus in environments where exploration capacity is limited [6, 7, 10, 11].

At a systems level, the framework interprets compositional generalization failures as manifestations of unbalanced trade-offs between data granularity and model abstraction. Fine-grained representations enhance local accuracy but may reduce robustness under combinatorial expansion, while coarse abstractions improve scalability at the cost of contextual sensitivity. Epistemically, the chosen abstraction level governs how resilient the system is to compositional novelty. Feedback structures within the framework can either reinforce existing biases—by repeatedly validating dominant representations—or alleviate them by exposing uncertainty boundaries and prompting reflective recalibration [1, 2, 5, 6]. The framework is shown in **Figure 1**.

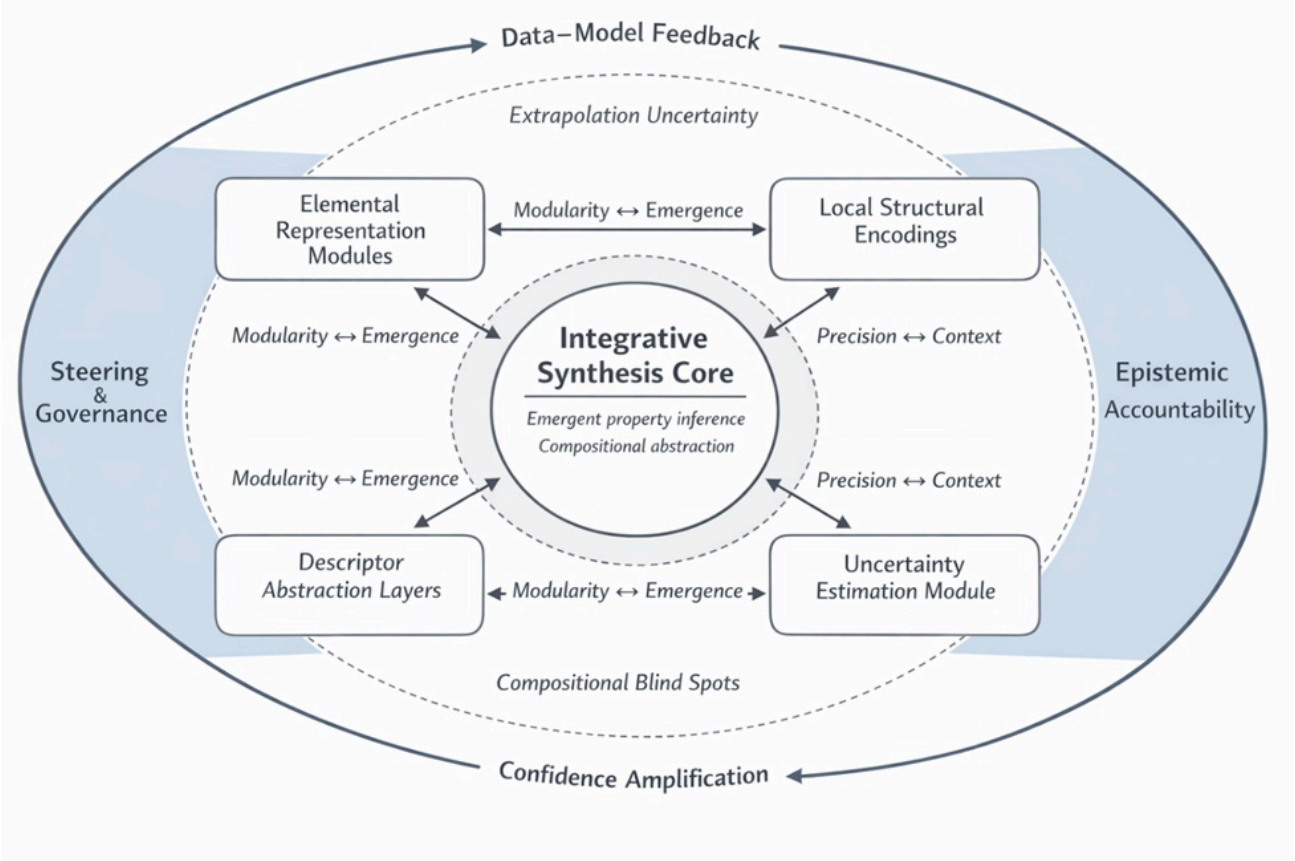

Figure 1. Conceptual framework for compositional generalization failure in materials AI

To situate compositional generalization within a broader interpretive landscape, **Table 3** maps its epistemic, systemic, and ethical dimensions, emphasizing how technical failures propagate through scientific decision-making.

Table 3. Epistemic, systemic, and ethical dimensions of compositional generalization failure

Dimension	Core issue	How failure manifests	Systemic feedback effect	Ethical concern raised
Epistemic	Incomplete representational invariance	Confident extrapolation beyond knowledge bounds	Reinforcement of false generality	Misinterpretation of model authority
Architectural	Modular bias toward local interactions	Weak synthesis of global composition	Narrow learning trajectories	Structural blind spots
Data ecosystem	Skewed compositional coverage	Interpolation-heavy performance	Self-confirming dataset expansion	Marginalization of rare materials
Pipeline dynamics	Model-guided exploration	Feedback amplification of bias	Reduced epistemic diversity	Path dependency in discovery
Governance	Performance-driven validation norms	Underscrutiny of extrapolative claims	Institutional overreliance on AI	Risk to sustainability and safety

Analytical implications

The conceptual framework outlined invites analytical implications that extend to the broader ecosystem of materials AI, interpreting compositional generalization failures as indicators of deeper systemic imbalances. In this view, the modular-feedback dynamics reveal how representational choices propagate through predictive pipelines, influencing the reliability of AI in compositional spaces [1, 5, 9]. For materials like glasses or composites, where property emergence defies linear combinations, these implications underscore trade-offs between computational tractability and interpretive depth, suggesting that overemphasis on modular efficiency may erode systemic coherence [4, 8, 13].

Epistemic reasoning further elucidates these implications by framing failures as opportunities to interrogate the boundaries of knowledge representation. Interaction dynamics between atomic descriptors and compositional contexts highlight feedback structures that can either amplify or dampen uncertainties, particularly in extrapolative scenarios [2, 6, 10]. Analytically, this implies a need to consider steering mechanisms that prioritize contextual integration, in which ethical considerations weigh the risks of deploying models in applications that demand high fidelity, such as battery materials optimization [1, 5, 9].

Systems-level insights integrate these elements, portraying compositional challenges as interwoven with data curation practices. The implications here are that biases in training sets perpetuate cycles of limited generalization, in which feedback from model performance informs future data collection, potentially creating virtuous or vicious loops [3, 7, 11]. In high-entropy materials, for example, analytical lenses reveal how compositional novelty exposes vulnerabilities in invariance assumptions, prompting reflections on the balance between innovation acceleration and epistemic caution [2, 6, 10].

Moreover, the framework's emphasis on trade-offs extends to interdisciplinary integrations, implying that insights from cognitive science—such as analogical reasoning—could inform materials AI by enhancing combinatorial flexibility without altering core architectures [4, 8, 12]. Ethical ramifications emerge in this analysis, interpreting failures as signals for responsible AI governance, ensuring that systemic feedback does not exacerbate inequalities in access to advanced materials technologies [3, 7, 11]. Ultimately, these implications foster a nuanced understanding of how compositional generalization shapes the trajectory of AI-driven materials science, advocating for interpretive strategies that harmonize modular strengths with holistic oversight [1, 2, 12, 13].

Results and Discussion

Delving deeper into the discourse, the conceptual interpretations of compositional generalization as a failure mode illuminate persistent tensions within materials AI paradigms. The interaction dynamics posited in the framework suggest that while AI excels in pattern recognition from fixed compositions, the recombination of elements introduces epistemic complexities that challenge systemic stability [3, 5, 9, 13]. This discussion integrates literature perspectives to explore how such dynamics manifest in practical domains, such as polymer design, where contextual dependencies override individual monomer predictions, revealing steering logics that favor adaptive rather than rigid learning [4, 8, 12].

Ethical reasoning threads through this analysis, interpreting the deployment of AI in materials discovery as a double-edged sword: accelerating progress while risking misaligned outcomes if compositional blind spots are ignored [1, 5, 6, 10]. Trade-offs in resource allocation become apparent, where the pursuit of broad compositional coverage competes with depth in specific systems, prompting systems-level insights into feedback structures that could realign priorities toward sustainable innovations [5, 6, 9, 13].

Furthermore, the synthesis of recent advancements highlights how architectural evolutions, such as topology-informed models, interact with compositional demands, offering analytical lenses into robustness enhancements [1, 2, 7, 11]. Yet, this integration cautions against overreliance on such evolutions, interpreting them as partial mitigations that still grapple with emergent behaviors in complex materials [1, 6, 10]. Epistemic considerations extend to the role of human-AI

collaboration, where interpretive frameworks could bridge gaps by leveraging expert intuition to navigate compositional uncertainties [2, 3, 8, 12].

In broader contexts, the discussion posits that compositional failures reflect macro-level shifts in scientific inquiry, where data-driven approaches intersect with theoretical foundations [3, 4, 7, 11]. This interplay underscores the need for conceptual recalibrations that view AI not as a panacea but as a component in a larger epistemic ecosystem, fostering resilience against failure modes through ongoing interpretive dialogue [2-6].

Conclusion

In synthesizing the conceptual contours of compositional generalization as a distinct failure mode in materials AI, this manuscript underscores the interpretive value of viewing these challenges through the lenses of modular interactions, systemic feedback, and epistemic trade-offs. The analytical implications and discussions reveal how these elements converge to shape the reliability and ethical deployment of AI in materials science, advocating for integrative strategies that enhance understanding without empirical overreach. Ultimately, this inquiry contributes to a scholarly appreciation of AI's transformative potential, tempered by a commitment to conceptual rigor in addressing its inherent limitations.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 22 Nov 2021 Revised: 23 Feb 2022 Accepted: 24 Apr 2022
Published online: 18 July 2022

References

Li K, Rubungo AN, Lei X, Persaud D, Choudhary K, DeCost B, et al. Probing out-of-distribution generalization in machine learning for materials. Commun Mater. 2025;6(1):9.
https://doi.org/10.1038/s43246-024-00731-w.

Shi X, Zhou L, Huang Y, Wu Y, Hong Z. A review on the applications of graph neural networks in materials science at the atomic scale. Mater Genome Eng Adv. 2024;2(2):e50.
https://doi.org/10.1002/mgea.50.

Peivaste I, Belouettar S, Mercuri F, Fantuzzi N, Dehghani H, Ghadiri M, et al. Artificial intelligence in materials science and engineering: Current landscape, key challenges, and future trajectories. Compos Struct. 2025;372(C):119419.
https://doi.org/10.1016/j.compstruct.2025.119419.

Shimakawa H, Kumada A, Sato M. Extrapolative prediction of small-data molecular property using quantum mechanics-assisted machine learning. Npj Comput Mater. 2024;10(1):11.
https://doi.org/10.1038/s41524-023-01194-2.

Zhang L, Banko L, Schuhmann W, Ludwig A, Stricker M. Composition-property extrapolation for compositionally complex solid solutions based on word embeddings. Digit Discov. 2025;4(1):1578-90.
https://doi.org/10.1039/D5DD00169B.

Li Q, Fu N, Omee SS, Hu J. Md-hit: Machine learning for material property prediction with dataset redundancy control. Npj Comput Mater. 2024;10(1):245.
https://doi.org/10.1038/s41524-024-01426-z.

Oliynyk A, Lee S, Jaffal E, Kumar NB, Kleinke H, Shiryaev D, et al. Composition and structure analyzer/featurizer for explainable machine-learning models to predict solid state structures. Digit Discov. 2025;4(4):548-60.
https://doi.org/10.1039/D4DD00332B.

Yang K, Song Y, Li Y, Smedskjaer MM, Bauchy M, Rosner F. Enabling extrapolation of young's modulus of cao-al2o3-sio2 ternary glasses by topology-informed machine learning. J Non Cryst Solids. 2025;666(C):123610.
https://doi.org/10.1016/j.jnoncrysol.2025.123610.

Gao L, Lin J, Wang L, Du L. Machine learning-assisted design of advanced polymeric materials. Acc Mater Res. 2024;5(5):571-84.
https://doi.org/10.1021/accountsmr.3c00288.

Wang Y, Li S, Li S, Li S, Chen M. Prediction, interpretation and extrapolation for shear modulus and bulk modulus of solid-state electrolytes based on machine learning. Mater Today Commun. 2024;38:108294.
https://doi.org/10.1016/j.mtcomm.2024.108294.

Cao Y, Taghvaie Nakhjiri A, Ghadiri M, Ghadiri M. Different applications of machine learning approaches in materials science and engineering: Comprehensive review. Eng Appl Artif Intell. 2024;135(C):108783.
https://doi.org/10.1016/j.engappai.2024.108783.

Hassan M, Fateh AA, Lin J, Zhuang Y, Lin G, Xiong H, et al. Unfolding explainable ai for brain tumor segmentation. Neurocomputing. 2024;599:128058.
https://doi.org/10.1016/j.neucom.2024.128058.

Alosious S, Jiang M, Luo T. Computation and machine learning for materials: Past, present, and future perspectives. MRS Bull. 2025;50(10):1212-24.
https://doi.org/10.1007/s43577-025-00959-y.