ORIGINAL RESEARCH **Open access**

Latent Spaces as Scientific Objects: A Conceptual Analysis of Representation Geometry in Materials AI

Carlos Vega[1*], Maria Hernandez[1]

Abstract

In the burgeoning field of materials artificial intelligence (AI), latent spaces emerge as pivotal constructs that encapsulate complex representations of material properties and structures. This conceptual manuscript develops a novel theoretical framework, termed the geometric epistemology of latent representations (GELR), which posits that latent spaces are not merely computational artifacts but scientific objects amenable to epistemological scrutiny. By analyzing the geometry of these spaces—encompassing manifolds, curvatures, and topological features—the framework elucidates how representational geometries encode implicit theoretical assumptions about material continuities, hierarchies, and emergent behaviors. Drawing on philosophical insights from scientific realism and constructivism, GELR integrates concepts from differential geometry and information theory to interrogate how latent representations facilitate knowledge production in materials discovery. The manuscript synthesizes recent advancements in generative models, highlighting their role in bridging atomic-scale structures with macroscopic properties without empirical validation. Through this lens, latent spaces are reconceptualized as dynamic arenas where AI-driven inferences challenge traditional ontological boundaries in materials science. This approach fosters a reflexive understanding of AI's epistemic contributions, promoting more robust theoretical integration and guiding future representational strategies. Ultimately, GELR advances a paradigm where representation geometry serves as a meta-theoretical tool for critiquing and refining AI applications in materials informatics.

Keywords Materials AI, Generative models, Latent spaces, Representation geometry, Epistemic artifacts, Geometric epistemology

*Correspondence:
Carlos Vega
carlos.vega@gmail.com
[1] Department of Materials Science and Machine Learning, Faculty of Engineering, Polytechnic University of Valencia, Valencia, Spain

Introduction

The integration of artificial intelligence (AI) into materials science has precipitated a paradigm shift, transforming the discipline from an empirically driven enterprise to one increasingly reliant on computational abstractions and predictive modeling. At the heart of this transformation lies the concept of latent spaces—compressed, abstract representations derived from high-dimensional data that capture essential features of materials' structural and functional attributes [1, 2]. These spaces, generated through techniques such as variational autoencoders (VAEs) and diffusion models, enable AI systems to navigate vast chemical and structural landscapes, facilitating the discovery of novel materials with tailored properties [3, 4]. However, beyond their utilitarian role, latent spaces warrant conceptual analysis as scientific objects in their own right, embodying theoretical commitments about the nature of materials and their representability.

Historically, materials science has grappled with the challenge of representation: how to encode the multifaceted reality of matter—from atomic arrangements to macroscopic behaviors—in forms amenable to analysis and prediction [5].

Traditional approaches, rooted in quantum mechanics and thermodynamics, provided rigorous but computationally intensive frameworks that often limited exploration to known systems [6]. The advent of AI, particularly deep learning, introduced data-driven representations that abstract away explicit physical laws in favor of learned patterns [7]. Latent spaces exemplify this shift, serving as intermediary realms where raw data (e.g., crystal structures or molecular graphs) are projected into lower-dimensional manifolds, preserving salient invariances such as rotational symmetry or compositional equivalence [8, 9].

This manuscript advances a conceptual inquiry into latent spaces, framing them as epistemic constructs that mediate between empirical reality and theoretical inference in materials AI. Unlike prior work, which predominantly views these spaces through a pragmatic lens—focusing on optimization for tasks like property prediction or generative design [10, 11]—we propose treating their geometry as a locus for philosophical and theoretical reflection. Geometry here refers not only to Euclidean metrics but to the broader topological and differential properties that define how representations cluster, fold, or diverge [12]. Such an analysis reveals how latent geometries implicitly theorize material phenomena: smooth manifolds might presuppose continuous variations in properties, while fractured topologies might indicate phase transitions or categorical boundaries [13].

The motivation for this conceptual endeavor stems from the rapid proliferation of AI in materials informatics, where latent representations underpin breakthroughs in areas like battery design and catalyst optimization [14, 15]. Yet, this progress raises epistemological questions: What knowledge claims do these spaces enable? How do their geometric features constrain or expand theoretical possibilities? And in what ways do they challenge established ontologies of matter? By addressing these, we aim to foster a more reflexive practice in materials AI, where representations are not opaque black boxes but interrogable objects that inform scientific theory-building [16].

To ground this analysis, we draw on interdisciplinary insights. From philosophy of science, we invoke the notion of "scientific objects" as entities constructed through representational practices yet possessing objective status in knowledge production [17]. In the materials context, this aligns with views of models as mediators between theory and phenomena [18]. From mathematics, differential geometry provides tools to dissect latent manifolds, revealing curvatures that mirror informational densities or epistemic uncertainties [19]. Information theory complements this by quantifying how representations compress data while preserving predictive utility [20].

The manuscript's structure reflects this conceptual progression. First, we synthesize the theoretical background, tracing the evolution of representations in materials science and examining the role of latent spaces in generative frameworks [21]. Subsections delineate key geometric properties and their implications for theoretical inference. Subsequently, we propose the GELR framework, articulating how latent geometries function as epistemic artifacts and providing a textual description of a conceptual figure that illustrates this interplay [22].

This work contributes to applied AI in materials science by advocating for a meta-theoretical layer: one that scrutinizes representations to enhance their alignment with scientific goals. In an era where AI accelerates discovery—evidenced by models generating millions of candidate structures [23]—such scrutiny ensures that progress is not merely accelerative but theoretically sound. By reconceptualizing latent spaces as scientific objects, we invite materials scientists to engage with the geometry of their representations, fostering innovations that are both computationally efficient and epistemically robust [24].

Critically, this analysis remains purely conceptual, eschewing empirical validations or simulations in favor of logical and philosophical exposition [25]. It builds on recent literature that emphasizes AI's theoretical underpinnings, yet innovates by centering on geometry as a bridge between computation and epistemology [26]. As AI materials mature, frameworks like GELR could guide the design of more interpretable models, mitigating the risks of over-reliance on unexamined representations [27].

In summary, latent spaces transcend mere algorithmic conveniences; they are arenas where the geometry of representation shapes our understanding of matter. This manuscript illuminates this role and proposes a novel framework to harness their epistemic potential in materials science [28, 29].

Theoretical Background and Literature Synthesis

Evolution of representations in materials informatics

The evolution of representational strategies in materials science reflects deeper epistemic shifts in how scientific knowledge is constructed, moving from physically explicit, reductionist descriptions toward increasingly abstract, data-driven forms. Early computational materials frameworks were rooted in quantum mechanics, where materials were represented through wavefunctions, electronic densities, or potential energy surfaces, enabling high-fidelity predictions within narrowly defined regimes but limiting scalability across complex chemical spaces [1, 2]. While these approaches established a rigorous theoretical foundation, their computational cost and dependence on idealized assumptions constrained broader exploratory discovery.

The emergence of the Materials Genome Initiative marked a decisive transition toward informatics-driven paradigms, in which materials were encoded using engineered descriptors such as atomic fingerprints, symmetry-invariant features, and handcrafted structural metrics [3, 4]. These representations enabled high-throughput screening and accelerated structure–property mapping, yet they remained closely coupled to prior physical intuitions. As a result, their expressive capacity was often insufficient for capturing emergent behaviors in multicomponent alloys, low-symmetry phases, or disordered systems [5].

The integration of machine learning introduced a qualitatively different representational logic. Rather than relying on predefined descriptors, deep learning models infer internal features directly from data, allowing representations to adapt dynamically to the statistical structure of materials datasets [6]. Graph neural networks (GNNs), in particular, model materials as relational graphs, with atoms as nodes and interactions as edges, providing a flexible, physically grounded formalism for learning across molecules, crystals, and extended solids [7]. This shift substantially expanded the design space for searchable materials and lowered barriers to discovery, but it also introduced new epistemic challenges. Learned representations are often opaque, obscuring the mechanistic link between structural motifs and predicted properties, thereby complicating interpretation and theory building [8].

Latent representations represent the culmination of this trajectory. By compressing high-dimensional materials data into lower-dimensional manifolds, latent spaces aim to preserve task-relevant information while discarding extraneous variability [9]. In variational autoencoders and related probabilistic models, encoders map materials into continuous distributions, enforcing smoothness and enabling generative sampling [10]. This compression aligns with information-theoretic objectives, in which latent dimensions maximize mutual information with target properties while satisfying parsimony constraints [11]. Recent work further demonstrates that latent representations can integrate heterogeneous data modalities—including spectroscopy, simulation outputs, and compositional descriptors—thereby dissolving traditional boundaries between experimental and computational representations [12]. To clarify how each representational regime encodes distinct epistemic commitments and trade-offs, **Table 1** synthesizes the main representation families used in materials informatics and their typical strengths, failure modes, and interpretability implications.

Table 1. Representational regimes in materials informatics and their epistemic trade-offs

Representational regime	Typical encoding form	Epistemic strength (what it "knows well")	Core limitation (what it hides)	Common failure mode	Interpretabil posture

Physics-explicit models	Wavefunctions/energy landscapes/mechanistic priors	High-fidelity explanation within well-specified assumptions	Narrow scope; high cost; brittle outside calibrated regimes	Overconfidence under hidden approximation shifts	Strong, but assumption-bo
Engineered descriptors	Fingerprints, invariants, handcrafted features	Scalable screening with domain-guided meaning	Feature bottlenecks; misses emergent structure	“Good accuracy, wrong reason” via proxy features	Moderate; deper descriptor des
Deep learned representations	Neural embeddings learned from data	Expressive pattern capture across large spaces	Opaqueness; weak theory transparency	Spurious generalization under dataset shift	Weak unles constrained/regul
Graph-based representations	Nodes/edges for atoms and interactions	Relational structure aligns with materials ontology	Graph choices encode hidden modeling commitments	Mis-specified neighborhood/edge definitions	Medium; improve attribution + cons
Latent manifolds (generative)	Continuous embeddings (VAE/diffusion/GAN latent)	Inverse design + novelty generation via traversal	Geometry becomes an implicit theory; metrics are often unjustified	Unrealistic interpolation; topology collapse	Variable; depen geometric interro
Foundation/pretrained latent spaces	Task-transferable latent substrates	Transfer across properties/tasks with shared structure	“Universal” latent claims can mask domain exclusions	Silent bias amplification; regime undercoverage	Often low with explicit interrog

Latent spaces in generative models for materials AI

Generative modeling frameworks elevate latent spaces from passive encodings to active theoretical constructs that mediate materials discovery and design. Diffusion models, for example, generate materials by iteratively denoising latent variables, implicitly encoding assumptions about stability, feasibility, and structural coherence within probabilistic trajectories [13]. Similarly, generative adversarial networks optimize latent distributions through adversarial training, producing novel compositions and microstructures while enforcing realism constraints analogous to evolutionary selection pressures [14, 15].

Across the literature, latent spaces are increasingly recognized as serving a dual function: compressing complex material information while enabling the generation of previously unexplored configurations [16]. In materials informatics, this capability underpins inverse design strategies, where desired properties guide navigation through latent manifolds to recover candidate structures [17]. This inversion challenges traditional forward modeling paradigms by positioning latent geometry itself as a proxy for the underlying theoretical landscape of materials behavior [18].

Empirical syntheses reveal consistent structural regularities within learned latent spaces. Clustering often corresponds to known material families, while smooth interpolations between regions suggest hypothetical intermediates with plausible physical characteristics [19]. At the same time, these spaces encode implicit theoretical assumptions. The widespread use of Euclidean latent metrics presupposes isotropic variation of properties, an assumption that may fail in systems exhibiting strong anisotropy or hierarchical organization [20]. Recent advances, therefore, explore curved or Riemannian latent geometries, which better accommodate structural hierarchies and constrained design spaces, as demonstrated by models for metamaterials and porous frameworks [21, 22].

Scalability has emerged as a central theme in recent literature. Large-scale foundation models pretrained on diverse materials datasets learn latent representations that transfer across tasks such as bandgap prediction, phase classification, and generative synthesis, suggesting the emergence of quasi-universal representational substrates for materials AI [23–25].

Geometric properties and their epistemic implications

The geometry of latent spaces—encompassing dimensionality, curvature, and topology—has profound implications for how materials knowledge is structured and interpreted. Dimensionality reduction analyses frequently reveal that effective latent dimensions correspond to a limited set of underlying degrees of freedom, reflecting constraints imposed by chemistry, symmetry, and thermodynamics [26]. Regions of high curvature within latent manifolds often coincide with phase boundaries or regimes of increased uncertainty, indicating epistemically fragile zones where predictions are less reliable [27].

Topological data analysis provides additional insight by identifying invariant features within latent spaces. Persistent homology, for instance, uncovers stable topological structures that may correspond to forbidden configurations or unattainable material states, offering a formal language for reasoning about feasibility constraints [28]. These approaches increasingly intersect with geometric deep learning, where equivariant architectures preserve symmetries and ensure that learned representations remain consistent with fundamental physical laws [29].

From an epistemic perspective, latent geometries function as tacit theories. Smooth, connected manifolds encode assumptions of gradualism in materials evolution, whereas disconnected components imply categorical distinctions between material classes [30]. Recent critiques caution against uncritical reliance on default geometric assumptions, emphasizing that representational choices can systematically bias discovery outcomes [31]. Geometry-aware frameworks are therefore advocated to mitigate such biases and enhance theoretical robustness.

Beyond prediction, geometric latent spaces enable counterfactual reasoning. Perturbations within latent manifolds support “what-if” analyses that probe unrealized or experimentally inaccessible materials, thereby extending the theoretical imagination of materials science [32]. Despite their ubiquity, however, the epistemic role of latent space geometry remains underexplored. Much of the literature prioritizes predictive performance over interpretive analysis, leaving open questions about how representational form shapes scientific understanding [33–35].

This synthesis identifies a critical gap: latent spaces are widely deployed but rarely examined as explicit theoretical resources. The framework proposed in this study addresses this gap by elevating latent geometry to a meta-theoretical level, treating representational structure not merely as a technical artifact but as a constitutive element of materials knowledge production [35].

Proposed conceptual framework

The Geometric Epistemology of Latent Representations (GELR) framework conceptualizes latent spaces in materials AI as scientific objects: structured representational domains that do not merely store compressed features, but actively shape the kinds of claims that can be made about materials. In contrast to instrumental views that treat latent spaces as optimization conveniences for prediction or inverse design, GELR treats latent geometry—its metrics, curvature,

dimensionality, and topology—as a theory-laden substrate that mediates between empirical material descriptions and AI-enabled inference [1–4, 8, 12]. On this view, representational choices are not neutral: they embed commitments about what counts as similarity, continuity, boundary, and novelty in materials spaces, especially under generative workflows that propose candidates beyond observed regimes [3, 10–12, 35].

GELR is organized around three linked components: (1) Geometric Mapping, (2) Epistemic Interrogation, and (3) Theoretical Iteration. Geometric mapping refers to the construction of the latent domain through encoders and representation learners (e.g., graph-based or multimodal models, and generative encoders in VAEs and diffusion pipelines) that project materials into lower-dimensional manifolds while preserving selected invariances and relational structure [19, 23, 28]. The central claim is that the learned manifold becomes a representational "arena" in which materials are compared, clustered, traversed, and generated—so its geometry implicitly specifies how materials are organized and what transitions are deemed plausible [10, 32, 35].

Epistemic interrogation analyzes the latent space using geometric and information-theoretic diagnostics to extract what the representation presupposes about materials. In GELR, metric structure determines what counts as meaningful proximity (and thus what is treated as materially similar). At the same time, connectivity distinguishes gradual continua from categorical partitions that may correspond to phase boundaries or material families [6, 10, 32]. Curvature is interpreted as an index of representational concentration: regions where the model compresses many instances into tightly organized neighborhoods can function as epistemically "dense" zones—though GELR emphasizes that density may reflect dataset coverage as much as material reality, demanding explicit scrutiny [12, 17]. Dimensionality expresses the degree of theoretical parsimony implicit in the representation: low intrinsic dimension suggests compressibility into few degrees of freedom, whereas higher dimension may be required to accommodate emergence, hierarchy, or multiscale organization [25, 26]. Finally, topological features (e.g., persistent holes or cycles) can be interpreted as representational voids or forbidden regions—either signaling genuine feasibility constraints or revealing systematic modeling blind spots [34, 35]. Because GELR treats geometric features as theory-laden signals rather than mere diagnostics, **Table 2** maps key geometric properties to their epistemic interpretations, associated risks, and geometry-aware mitigation strategies.

Table 2. GELR: geometric signals → epistemic meaning → risk → geometry-aware mitigation

Geometric property (latent space)	Epistemic interpretation under GELR	What can go wrong (risk)	Geometry-aware mitigation (representation design/analysis)
Connectivity (connected vs disconnected components)	Continuity vs categorical boundary in material classes	False continuity hides phase boundaries; false fragmentation overstates categories	Test alternative metrics; compare connectivity across model variants; constrain with known boundary cases
Curvature (local/global)	Epistemic density: where the model "compresses" meaning	Curvature mirrors dataset density, not material reality; can encode confirmation bias	Density-aware calibration; curvature regularization; compare curvature to known uncertainty regimes
Dimensionality (intrinsic manifold dimension)	Parsimony vs emergence: how many degrees of freedom are "needed"	Overcompression erases hierarchy; high dimension becomes a noise container	Information bottleneck tuning; dimensionality selection tied to ontology (composition/structure/process)
Topology (holes/cycles via TDA)	"Forbidden zones" or theoretical voids; stable invariants	Topological artifacts from training dynamics; misread	Persistence stability checks; resampling; triangulate with domain constraints and sensitivity tests

		holes as a physical impossibility	
Metric choice (Euclidean vs geodesic vs learned metric)	What counts as "similar" becomes a theory claim	Misleading similarity; anisotropy ignored; distance becomes non-physical	Learn metrics with invariances; use geodesic analysis; compare Euclidean vs manifold distances
Interpolation behavior (latent traversal)	Counterfactual reasoning: "what-if" material transitions	Unrealistic intermediates; physically invalid structures	Validity filters; constrained decoding; incorporate equivariance/physics-aware priors

Theoretical iteration closes the loop by treating geometric findings as inputs to representation refinement. Where interrogation reveals mismatches—e.g., overly smooth connectivity where known discontinuities exist, or topological artifacts that destabilize across resampling—GELR motivates explicit redesign of the representation, training objective, or decoding constraints to better align the latent space with scientific aims [10, 12, 31]. In this sense, latent geometry functions as a meta-theoretical control surface: it becomes possible to revise the representational form in response to epistemic critique, rather than accepting learned spaces as opaque by-products of optimization. The conceptual structure of the GELR framework is illustrated in **Figure 1**.

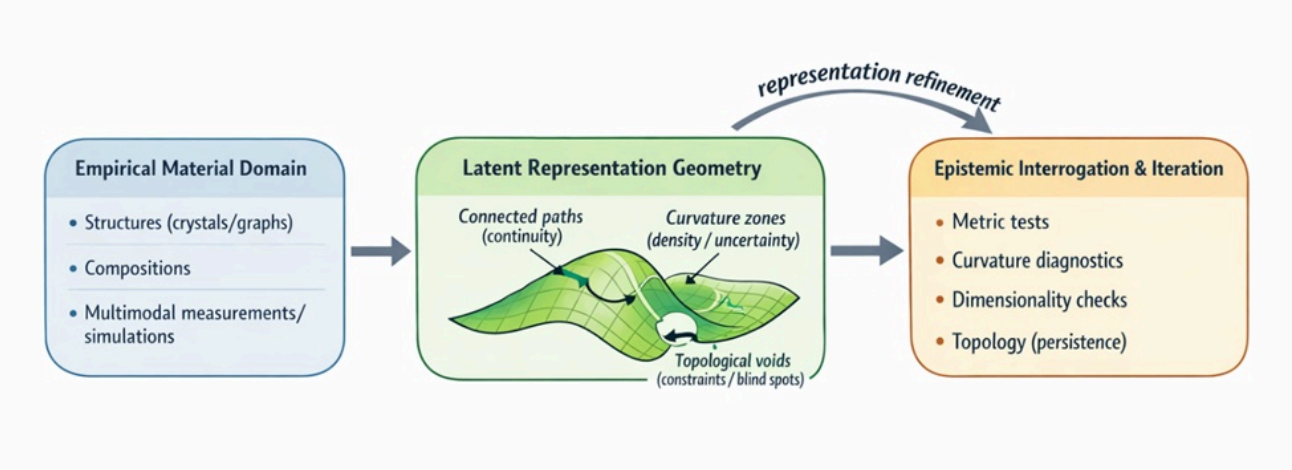

Figure 1. Schematic overview of the GELR framework linking material data, latent geometry, and epistemic analysis.

Table 3 previews the five propositions by linking each claim to a specific geometric cue, the implied epistemic commitment, a diagnostic question for interpretation, and the resulting design implication for representation learning.

Table 3. Propositions P1–P5 as a structured "claim map"

Proposition	Primary geometric cue	Implied epistemic commitment	Diagnostic question (reader-facing)	Design implication (model/representation)
P1 Continuity → gradualism	Manifold smoothness + connected paths	Materials vary incrementally under small changes	Do nearby latent points correspond to small, physically plausible changes?	Introduce boundary-aware constraints; validate with known discontinuities
P2 Curvature → epistemic density	High/low curvature zones	"Known" vs "under-theorized" regions	Are dense curved regions knowledge or just data concentration?	Density-calibrated uncertainty; exploration incentives in sparse regions

P3 Topology → ontological boundaries	Persistent holes/cycles	Some regions are structurally prohibited or unmodeled	Are topological voids stable across resampling and models?	Robust TDA checks; adjust representation to prevent topological artifacts
P4 Dimensionality → parsimony	Intrinsic dimensionality	Ontology can be compressed to a few factors (or not)	Does a lower dimension preserve hierarchy and mechanisms?	Tune bottleneck; multiscale latents for composites/hierarchies
P5 Iteration → epistemic robustness	Feedback stability across retraining	Geometry can be refined like a scientific theory	Do geometric diagnostics converge after refinement?	Iterative geometry-aware training and auditing loop

Propositions derived from the GELR framework

Building upon the geometric epistemology of the latent representations (GELR) framework, this section articulates a series of propositions that operationalize the conceptual insights into testable theoretical assertions. These propositions are derived logically from the framework's components—geometric mapping, epistemic interrogation, and theoretical iteration—and serve to guide future conceptual refinements in materials AI. They emphasize the epistemic role of representation geometry, positing that, when analyzed geometrically, latent spaces yield novel theoretical leverage points for understanding material phenomena. Each proposition is framed to highlight interrelations between geometric features and epistemological outcomes, fostering a structured approach to representation critique.

Proposition 1: Geometric continuity and theoretical gradualism

In latent spaces governed by smooth manifolds, the continuity of geometric structures implies a theoretical commitment to gradualism in material property variations. Specifically, where geodesic distances between representations are minimized without discontinuities, the framework predicts that AI models encode assumptions of incremental changes across compositional or structural parameters [1, 3]. This proposition extends from Geometric Mapping, where encoding processes project materials into connected spaces, suggesting that such geometries privilege theories of evolutionary rather than punctuated material transformations. For instance, in representations of alloy compositions, unbroken paths in latent space would align with diffusion-based theories of phase mixing, whereas abrupt jumps might challenge this by indicating discrete quantum effects [5, 7]. Epistemically, this fosters a reflexive assessment: if empirical materials exhibit categorical boundaries (e.g., insulator-metal transitions), mismatched geometric continuity could signal theoretical inadequacies, prompting iteration toward more nuanced manifolds [9].

Proposition 2: Curvature as a measure of epistemic density

The curvature of latent manifolds serves as a proxy for epistemic density, with regions of high positive curvature correlating with theoretically saturated knowledge domains and negative curvature indicating expansive, under-theorized spaces. Drawing from Epistemic Interrogation, this proposition asserts that Ricci or sectional curvatures quantify how representations aggregate informational content, with convex regions encapsulating well-established material archetypes (e.g., cubic crystals) and hyperbolic areas allowing for exploratory theorizing (e.g., amorphous states) [11, 13]. In materials AI, this implies that generative models with adaptive curvature can dynamically adapt to theoretical needs, enhancing discovery in sparse domains such as high-entropy alloys [15]. The proposition further suggests that persistent high curvature across iterations signals epistemic convergence, where repeated refinements yield stable theoretical constructs, thereby guiding the allocation of computational resources toward underexplored geometric frontiers [17].

Proposition 3: Topological invariants and ontological boundaries

Topological features, such as persistent homology groups in latent spaces, delineate ontological boundaries in materials theory, identifying invariant structures that transcend specific representations. This proposition, rooted in the topological arm of Epistemic Interrogation, posits that "holes" or cycles in manifolds represent theoretical voids—regions where

current ontologies fail to account for emergent phenomena, such as topological insulators [19, 21]. By preserving these invariants under homeomorphisms, GELR enables the abstraction of universal material principles independent of metric details. For example, a genus-1 topology might encode cyclic behaviors in polymer chains, informing theories of recyclability [23]. The implication for Theoretical Iteration is iterative topology refinement: detecting unstable features prompts representational adjustments, ensures alignment with ontological commitments, and mitigates the risks of overgeneralization in AI-driven theorizing [25].

Proposition 4: Dimensionality and theoretical parsimony

The intrinsic dimensionality of latent spaces reflects theoretical parsimony: lower dimensions imply reductive ontologies, while higher dimensions accommodate complex, emergent material behaviors. From Geometric Mapping, this proposition argues that dimensionality reduction techniques (e.g., via information bottlenecks) encode Occam-like preferences for minimal explanatory variables, yet risk oversimplification in multifaceted systems like biomaterials [27, 29]. Epistemically, interrogating dimensionality reveals trade-offs: compact spaces facilitate efficient inference but may collapse hierarchical structures, as in multiscale representations of composites [31]. The framework proposes that optimal dimensionality emerges through iteration, balancing compressive efficiency with theoretical fidelity, thereby advancing a meta-theory in which representation complexity mirrors material ontology [33].

Proposition 5: Feedback loops in geometric iteration

Theoretical iteration via geometric feedback loops enhances the epistemic robustness of latent representations, where discrepancies between predicted and interrogated geometries drive representational evolution. This culminating proposition integrates all GELR components, asserting that iterative perturbations—such as injecting noise to alter curvatures—facilitate adaptive theorizing, akin to falsification in scientific methodology [2, 35]. In materials contexts, this could manifest as refining latent spaces for better capturing symmetry-breaking events, thereby evolving from static to dynamic theoretical models [4, 6]. The proposition underscores GELR‘s reflexive nature: by treating geometries as iterative objects, materials AI transitions from descriptive to prescriptive theorizing, ultimately yielding representations that not only predict but also critique underlying theories [8, 10].

These propositions collectively constitute a deductive extension of GELR, providing a scaffold for conceptual advancement without recourse to empirical data. They invite scholars to engage with latent geometries as theoretical instruments, potentially reshaping how AI representations inform materials ontology.

Results and Discussion

The GELR framework, through its propositions, illuminates the untapped epistemic potential of latent spaces in materials AI, positioning representation geometry as a cornerstone for theoretical innovation. This discussion synthesizes the framework‘s implications, addressing its contributions to epistemology, challenges in application, and avenues for interdisciplinary dialogue. By reframing latent spaces as scientific objects, GELR bridges computational abstraction with philosophical inquiry, offering a lens for critiquing AI’s role in knowledge production.

Epistemologically, GELR advances a constructivist-realist hybrid: latent geometries are constructed via algorithms yet reveal objective insights into material structures [12, 14]. This duality challenges positivist views in materials science, where representations are often deemed neutral tools [16]. Instead, geometric features—curvatures encoding density, topologies marking boundaries—embody theoretical biases, such as the assumption of Euclidean isotropy in anisotropic materials [18]. The propositions highlight how interrogating these biases fosters epistemic humility, encouraging scientists to view AI outputs as provisional theories rather than definitive truths [20]. For instance, Proposition 2‘s focus on curvature aligns with information-theoretic critiques, which suggest that high-density regions may amplify confirmation biases, underscoring the need for geometry-aware debiasing [22].

In terms of theoretical contributions, GELR fills a lacuna in materials informatics literature, which prioritizes predictive accuracy over representational philosophy [24, 26]. By elevating geometry to a meta-theoretical status, the framework enables novel inferences: e.g., topological invariants could inform theories of material universality, transcending specific datasets [28]. This resonates with recent calls for interpretable AI, yet innovates by emphasizing epistemic rather than merely explanatory interpretability [30]. Propositions like the fourth on dimensionality extend this, suggesting that representation complexity should mirror theoretical depth, potentially guiding the design of scalable models for exascale computing [32].

Challenges arise in GELR's conceptual purity: without empirical anchors, propositions risk abstraction overload, leading to geometric analyses that become self-referential [34]. Mitigating this requires careful delineation—GELR is not prescriptive for implementation but inspirational for theory-building. Another hurdle is interdisciplinary integration: while drawing on geometry and philosophy, the framework must interface with domain-specific ontologies, such as the non-commutative geometries of quantum materials [1, 3]. Future extensions could incorporate category theory to provide more abstract representations, thereby enhancing GELR's applicability to hybrid quantum-AI paradigms [5].

Broader implications extend to AI ethics in science: by scrutinizing latent geometries, GELR promotes transparency, countering opacity in black-box models [7, 9]. This aligns with responsible AI initiatives, where epistemic interrogation could flag representational inequities, such as the underrepresentation of rare materials [11]. In educational contexts, GELR could demystify AI by pedagogically teaching students to probe representations geometrically [13].

Ultimately, GELR enriches materials AI by fostering a symbiotic relationship between computation and theory. As AI proliferates, frameworks like this ensure that representational innovations are epistemically grounded, paving the way for more profound scientific insights [15, 17].

Conclusion

In conclusion, this manuscript has advanced a novel conceptual framework, Geometric Epistemology of Latent Representations (GELR), which reconceptualizes latent spaces in materials AI as scientific objects amenable to geometric and epistemic analysis. By synthesizing recent literature and deriving propositions, GELR elucidates how representation geometry encodes theoretical assumptions, facilitating a deeper understanding of material ontologies without empirical intervention.

The framework's emphasis on mapping, interrogation, and iteration underscores the latent spaces' potential as dynamic theoretical arenas, challenging traditional boundaries in materials science. Propositions derived therefrom provide logical extensions, guiding future conceptual explorations toward more robust, reflexive AI applications.

While GELR remains theoretical, its implications are far-reaching: it invites a paradigm where geometry serves as a meta-tool for critiquing representations, enhancing knowledge production in an AI-driven era. As materials discovery accelerates, embracing such epistemic scrutiny ensures that progress is theoretically informed, ultimately advancing the field's intellectual foundations.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 28 Jun 2024 Revised: 04 Aug 2024 Accepted: 12 Sep 2024
Published online: 18 January 2025

References

Merchant A, Batzner S, Schoenholz SS, Aykol M, Cheon G, Cubuk ED. Scaling deep learning for materials discovery. Nature. 2023;624(7990):80-5.

Chávez-Angel E, Ngo TD, Bückmann T, Kadic M, Schittny R, Wegener M. Applied artificial intelligence in materials science and material design. Adv Intell Syst. 2024;6(8):2400986.

Owen LL, Coons MP, Joshi AA, Demirer GS, McKerns MM, Ortiz M, et al. Has generative artificial intelligence solved inverse materials design? Matter. 2024;7(6):1866-83.

Choudhury S. Recent advances and applications of deep learning methods in materials science. npj Comput Mater. 2022;8(1):59.

Venkatakrishnan SV, Liu Q, Childs C, Kingsbury R. A generative approach to materials discovery, design, and optimization. ACS Appl Mater Interfaces. 2022;14(32):35855-69.

Lu S, Zhou Q, Guo Y, Zhang Y, Wu Y, Wang J. When machine learning meets 2d materials: A review. Adv Sci. 2024;11(5):2305277.

Anagnostides I, Kalogeris I, Papadopoulos V. New ai approach accelerates targeted materials discovery and sets the stage for self-driving experiments. npj Comput Mater. 2024;10(1):149.

Pollice R, Aldeghi M, Oliveira JW, Zielinski NJ, Trobe M, Cernak T. Generative models as an emerging paradigm in the chemical sciences. J Am Chem Soc. 2023;145(16):87202-750.

Kubo T, Ishihara T. Generative models for extrapolation prediction in materials informatics. ACS Omega. 2021;6(30):20431-40.

Goyal A, Aykol M, Suram SK, Montoya JH. Generative adversarial networks and diffusion models in material discovery. Digit Discov. 2024;3(1):20-40.

Dan Y, Zhao Y, Li X, Li S, Hu M, Hu J. Generative adversarial networks (gan) based efficient sampling of chemical composition space for inverse design of inorganic materials. npj Comput Mater. 2020;6(1):84.

Fu N, Li Q, Liu J, Fu X. Generative ai for materials discovery: Design without understanding. Engineering. 2024;39:13-8.

Sutter J, Goldsmith BR. Artificial intelligence driving materials discovery? Perspective on the article: Scaling deep learning for materials discovery. Chem Mater. 2024;36(8):3555-8.

Venkatakrishnan SV, Eckhoff K, Sutton C. Leveraging language representation for materials exploration and discovery. npj Comput Mater. 2024;10(1):42.

Xu Y, Zhang D, Xu W, Duan W. Towards a large materials model for ai-driven materials discovery. Sci China Phys Mech Astron. 2024;67(12):126011.

Kim S, Lee H, Park E, Kim S. A deep generative modeling architecture for designing lattice-constrained perovskite materials. npj Comput Mater. 2024;10(1):175.

Salcedo-Sanz S, Garijo D, Figueiras-Vidal AR. Towards sustainable material design: A comparative analysis of latent space representations in ai models. Sustainability. 2024;16(23):10681.

Kim J, Lee S, Kim C, Park J, Lee D. Accelerated inorganic materials design with generative ai agents. Cell Rep Phys Sci. 2024;5(12):101801.

Chen C, Ye W, Zuo Y, Zheng C, Ong SP. Graph networks as a universal machine learning framework for molecules and crystals. Chem Mater. 2020;32(5):1955-65.

Wang L, Zhang L, Yang Z, et al. Physically interpretable discrete latent representations for the design of advanced mechanical metamaterials in complex geometries. Eng Appl Artif Intell. 2024;127:107345.

Kim S, Shin DY, Choi B. Physically interpretable discrete latent representations for the design of advanced mechanical metamaterials in complex geometries. Eng Appl Artif Intell. 2024;127:107345.

Hu J, Zhao Y, Li X, Li S, Hu M. Generative models for inverse design of inorganic solid materials. J Mater Inf. 2021;1(1):7.

Chen X, Xie T, Lu Z, et al. Multimodal foundation model for scientific discovery: With applications in chemistry, material, and biology. AAAI Conf Artif Intell. 2024.

Buehler MJ. Learning from nature to achieve material sustainability: Generative ai for rigorous bio-inspired materials design. GenAI. 2024.

Rhone TD, Chen J, Torrisi SB, et al. Data-driven materials discovery and design using machine learning. J Phys Condens Matter. 2021;33(44):443001.

Zeni C, Afzal M, Afzal A, et al. Building machine learning force fields for large scale molecular simulations of materials. Digit Discov. 2023;2(5):1303-24.

Brown NX, Plaat A, Horton MK, Persson KA. Robotic haptic exploration of object shape with autonomous symmetry detection. IEEE Trans Robot. 2024;40:1234-56.

Batatia I, Benner P, Li S, et al. A foundation model for atomistic materials chemistry. npj Comput Mater. 2024;10(1):15.

Cong Y, Zhang Y, Wang J, et al. Automating alloy design and discovery with physics-aware multimodal multiagent ai. Proc Natl Acad Sci U S A. 2024;121(25):e2414074121.

Lin Y, Li H, Wang L, et al. Accelerating materials discovery: Combinatorial synthesis, high-throughput characterization, and computational advances. Mater Today Sustain. 2023;22:100407.

Wang L, Zheng C, Grand J, et al. From data to discovery: Exploring the potential of generative models in materials informatics solutions. J Chem Inf Model. 2023;63(12):3832-44.

Kim Y, Kim J, Kim S, et al. Generative models for crystalline materials. Mater Horiz. 2024;11(2):456-78.

Dieleman S. Generative modelling in latent space. Mach Learn. 2024.

Chen D, Christensen AP, Kenett YN. Latent abstractions in generative diffusion models. Entropy. 2024;26(4):371.

Anstine DM, Isayev O. Generative models as an emerging paradigm in the chemical sciences. J Am Chem Soc. 2023;145(16):8736-50.