ORIGINAL RESEARCH **Open access**

Inference Without Ground Truth: A Conceptual Theory of Validation in Materials AI

Michael Turner[1*], Daniel Brooks[1], Olivia Grant[2]

Abstract

In the domain of materials artificial intelligence (AI), the lack of reliable ground truth poses significant challenges for validating inferential processes. This conceptual manuscript develops a novel theoretical framework for understanding validation dynamics in contexts where empirical benchmarks are scarce or contested. Drawing on recent literature in materials informatics, data bias, and epistemic values in science, the framework interprets validation as an integrative system of interaction dynamics between AI-generated inferences and epistemic feedback structures. It explores the analytical implications of managing trade-offs between uncertainty and bias, emphasizing systems-level insights into how inferential reliability emerges from iterative conceptual interpretations rather than direct empirical confrontation. The framework highlights ethical reasoning in steering logics that govern data curation and model deployment in materials discovery. By synthesizing these elements, the paper offers interpretive tools for navigating the epistemic landscape of AI-driven materials science, fostering more robust conceptual integration without relying on propositional claims or empirical validation. This approach contributes to applied AI in materials by illuminating pathways for enhanced inferential integrity amid inherent data ambiguities.

Keywords Materials AI, Validation dynamics, Ground truth absence, Inference interpretation, Data bias trade-offs, Epistemic feedback

*Correspondence:
Michael Turner
michael.turner@gmail.com
[1] Department of Materials Informatics, Faculty of Engineering, University of Manchester, Manchester, United Kingdom
[2] Department of Artificial Intelligence Systems, Faculty of Computer Science, University of Birmingham, Birmingham, United Kingdom

Introduction

The integration of artificial intelligence (AI) into materials science has transformed the landscape of discovery and design, enabling unprecedented scales of computational exploration and predictive modeling [1, 2]. From accelerating the identification of novel compounds to optimizing properties in complex systems, AI tools have become indispensable in addressing the vast chemical space that traditional experimental methods struggle to traverse [3]. However, this advancement poses profound epistemic challenges, particularly in validating inferences when ground truth—defined as verifiable empirical evidence—is either unavailable, incomplete, or inherently uncertain [4]. In materials AI, ground truth often relies on experimental data that is costly, time-consuming to obtain, or subject to variability due to measurement conditions, leading to scenarios where inferences must be assessed through alternative conceptual lenses [5].

This manuscript delves into these challenges by developing a conceptual theory of validation tailored to materials AI, focusing on inference processes that operate without direct access to unequivocal ground truth. Rather than proposing empirical solutions or testable models, the work emphasizes analytical implications and interaction dynamics that shape the reliability of AI outputs in materials contexts [6]. It interprets validation not as a static checkpoint but as a fluid system

of epistemic interactions, where biases in data and models interact with scientific values to influence inferential outcomes [7]. Such an approach is timely, as the proliferation of machine learning techniques in materials informatics has amplified concerns over data quality, model generalizability, and the ethical dimensions of automated discovery [8, 9].

The motivation for this conceptual exploration stems from the recognition that materials science increasingly relies on AI for high-throughput screening and generative design, where the sheer volume of predictions outpaces experimental verification [10]. For instance, in alloy design or polymer engineering, AI models generate hypotheses about material properties based on sparse datasets, raising questions about how to discern meaningful insights from artifacts of algorithmic bias [11]. Traditional validation paradigms, rooted in statistical metrics such as accuracy or precision, evaluated against labeled data, falter in these settings because the “truth” is often provisional or context-dependent [12]. This gap underscores the need for a theoretical framework that integrates epistemic reasoning with systems-level insights, allowing researchers to navigate trade-offs in uncertainty without defaulting to empirical resolutions [13].

Literature in the field reveals a growing discourse on these issues. Recent works have highlighted how data biases—arising from historical experimental preferences or database incompleteness—permeate AI models, potentially skewing inferences toward overrepresented material classes [14]. Concurrently, discussions on epistemic values in science emphasize the role of human judgment in interpreting AI outputs, suggesting that validation involves more than technical metrics; it encompasses ethical considerations of trustworthiness and societal impact [15]. By synthesizing these strands, this paper advances a novel interpretive lens, viewing validation as an emergent property of feedback structures between inference engines and conceptual safeguards [16].

The structure of the manuscript proceeds as follows. Following this introduction, the theoretical background synthesizes key literature on materials AI, ground truth challenges, data bias, and epistemic values, organized under subheadings to facilitate integrative analysis [17]. Subsequently, the proposed conceptual framework is articulated, detailing its core dynamics and including a textual description of **Figure 1**, which illustrates the interaction pathways [18]. The manuscript adheres to a purely conceptual orientation, avoiding any empirical elements, and draws exclusively on peer-reviewed sources to ensure contemporaneity [19].

In essence, this work contributes to applied AI in materials science by offering tools for conceptual interpretation that enhance the epistemic robustness of inferences. It invites scholars to reconsider validation as a multifaceted process of steering logics and trade-offs, ultimately fostering more nuanced engagements with AI in the pursuit of materials innovation [20].

Theoretical Background and Literature Synthesis

The theoretical foundations of validation in materials artificial intelligence (AI) emerge from the intersection of materials informatics, philosophy of science, and systems theory. Unlike classical validation paradigms—where model outputs can be benchmarked against stable empirical ground truth—materials AI operates under conditions of epistemic incompleteness, data scarcity, and constructed reference standards. Recent literature increasingly treats validation not as a binary confirmation exercise but as a distributed inferential process shaped by data representations, model architectures, epistemic values, and institutional constraints. This section synthesizes scholarship, organizing it into four interrelated thematic domains that together illuminate how inferential credibility is established in the absence of definitive ground truth.

AI-Mediated materials discovery and design

Predictive inference in high-dimensional materials spaces

Recent advances in machine learning have reconfigured materials discovery by enabling predictive inference across vast, high-dimensional design spaces that exceed the tractability of conventional physics-based simulations [1–5]. Models such

as graph neural networks, variational autoencoders, and diffusion-based generators infer structure–property relationships by exploiting statistical regularities embedded in large, heterogeneous datasets [2, 4]. Rather than explicitly encoding physical laws, these approaches derive inferential leverage from representational capacity, enabling the identification of latent correlations inaccessible to manual or rule-based analysis [6, 7].

This shift reframes discovery as an inferential activity conducted within abstract representation spaces. Materials candidates are no longer validated solely through direct correspondence with experimental measurements but through consistency across learned patterns, surrogate objectives, and internal coherence within model ensembles [8, 9].

Systems-level inference and feedback dynamics

The literature increasingly conceptualizes materials AI systems as feedback-driven inferential ecosystems rather than isolated predictors [9, 10]. Model outputs influence subsequent data collection, screening priorities, and experimental allocation, creating recursive loops in which inference and evidence co-evolve. Validation, in this context, becomes distributed across iterative refinement cycles rather than localized at a single benchmarking step.

Without stable ground truth, inferential confidence is often established through proxy criteria, such as agreement with known physical constraints, robustness under perturbation, or convergence across independently trained models [11]. These criteria reflect a broader systems-level understanding of inference, where credibility arises from relational stability rather than direct verification.

Inverse design and amplified uncertainty

Inverse design frameworks further intensify validation challenges by reversing the inferential direction: desired properties guide the generation of candidate materials rather than emerging as predictions from known structures [12, 13]. While powerful, this paradigm compounds epistemic uncertainty, as errors propagate backward from abstract objectives into proposed material configurations. The literature emphasizes that inverse design outputs frequently lack empirical anchors, requiring validation to rely on conceptual plausibility, physical consistency, and comparative screening rather than direct confirmation.

The problem of ground truth in materials data

Constructed reference standards and surrogate truth

Ground truth in materials science is rarely absolute. Experimental measurements are mediated by instrumentation limits, synthesis variability, and theoretical assumptions, while computational databases rely on approximations whose validity depends on modeling choices [14–16]. Large repositories such as Materials Project, AFLOW, and related resources provide essential scaffolding for materials AI, yet function as surrogate truth systems rather than definitive references.

Recent literature reframes ground truth as an epistemic construct stabilized through community consensus rather than a fixed ontological anchor [17, 18]. Under this view, validation becomes an exercise in aligning AI inferences with accepted representational conventions rather than uncovering immutable facts.

Validation under throughput asymmetry

High-throughput AI pipelines routinely generate predictions at a pace that far exceeds the capacity for experimental validation [19]. This asymmetry forces prioritization decisions about which inferences warrant further scrutiny. As a result, validation is increasingly conceptualized as a selective and value-laden process, shaped by feasibility constraints, perceived importance, and downstream application contexts.

Trade-offs between discovery speed and epistemic reliability are prominent in the literature, particularly when simulated data substitutes for empirical measurement [20]. Rather than treating these trade-offs as technical shortcomings, recent

syntheses argue for explicit interpretive frameworks that acknowledge ground truth as provisional, layered, and context-dependent [21].

Data bias and inferential distortion in materials informatics

Structural biases in materials datasets

Materials datasets reflect historical research priorities, economic incentives, and experimental convenience, leading to systematic over-representation of certain chemistries, crystal structures, and property regimes [22–24]. AI models trained on such data inherit and often amplify these skews, producing inferences that appear robust within familiar regions while failing under extrapolation [25].

The literature emphasizes that these distortions are not merely statistical artifacts but epistemic consequences of uneven knowledge production. Validation practices that rely on held-out subsets of similarly biased data risk reinforcing false confidence.

Bias as a systems-level phenomenon

Recent work frames bias as an emergent property of interactions between data curation, model optimization, and institutional decision-making [26–28]. Efforts to diversify datasets introduce new uncertainties, as expanding chemical space often reduces data density and predictive precision. Consequently, validation involves navigating trade-offs between representational breadth and inferential stability [27].

This systems perspective positions bias mitigation as an ongoing interpretive activity rather than a one-time corrective step, further complicating validation criteria in materials AI.

Ethical dimensions of biased inference

Bias in materials AI carries ethical implications, influencing which materials are explored, which applications are prioritized, and which societal needs are addressed [29]. Validation, therefore, extends beyond technical performance to encompass questions of distributive impact and scientific responsibility. The literature increasingly argues that inferential credibility cannot be decoupled from the values embedded in data selection and modeling objectives.

Epistemic values and ethical reasoning in scientific AI

Epistemic values as validation criteria

In the absence of definitive ground truth, epistemic values—such as reliability, transparency, robustness, and coherence—serve as guiding principles for interpreting AI outputs [30–32]. Rather than functioning as abstract ideals, these values actively shape model design, evaluation practices, and acceptance thresholds within the materials AI community.

Validation thus becomes a value-conditioned process, where inferential legitimacy depends on alignment with shared norms about what constitutes trustworthy scientific knowledge.

Value-driven steering logics in AI workflows

Recent scholarship highlights how epistemic values are operationalized through steering logics embedded in AI workflows, including choices of loss functions, uncertainty metrics, and stopping criteria [33]. These design decisions implicitly prioritize certain forms of knowledge—such as predictive accuracy or physical plausibility—over others, shaping which inferences are deemed valid.

Systems-level analyses reveal that such values propagate through feedback loops, influencing not only model outputs but also subsequent data generation and experimental validation strategies [34, 35].

Integration of ethical and epistemic reasoning

The literature increasingly rejects sharp distinctions between epistemic and ethical evaluation, emphasizing their entanglement in scientific AI [34]. In materials contexts, validation involves assessing not only whether an inference is internally coherent but also whether its downstream implications align with broader scientific integrity and societal responsibility.

Synthesis and conceptual implications

Taken together, these thematic strands portray validation in materials AI as a dynamic, multi-layered inferential practice rather than a definitive test against ground truth. AI-mediated discovery, constructed reference standards, systemic data bias, and value-laden reasoning collectively shape how credibility is established. This synthesis sets the conceptual foundation for the framework developed in the subsequent section, which formalizes these interactions into a coherent model of validation dynamics under epistemic uncertainty.

Proposed conceptual framework

The proposed conceptual framework interprets validation in materials AI as an emergent system of interaction dynamics, where inferences are shaped by epistemic feedback structures in the absence of ground truth. This approach moves beyond traditional metrics to emphasize analytical implications, conceptual interpretations, and trade-offs that govern inferential reliability. At its core, the framework envisions validation as a cyclical process involving steering logics that navigate uncertainties through integrative reasoning.

Central to the framework is the recognition that inference is a multifaceted interaction between AI algorithms and material data ecosystems. Without ground truth, reliability emerges from the dynamics of bias detection and epistemic alignment, where trade-offs in data representation influence outcome interpretations [1, 14]. Systems-level insights reveal how feedback loops—connecting model outputs back to data curation—facilitate adaptive validation, enabling conceptual refinements over time [22, 26].

Ethical reasoning infuses the framework, positioning validation as an epistemic endeavor that balances innovation with cautionary trade-offs [30, 33]. For instance, interactions between biased datasets and inferential engines highlight the need for interpretive safeguards to mitigate amplification effects [23, 27]. The framework thus offers tools for analyzing these dynamics, fostering deeper understandings of how validation structures evolve in materials contexts, as shown in **Figure 1**.

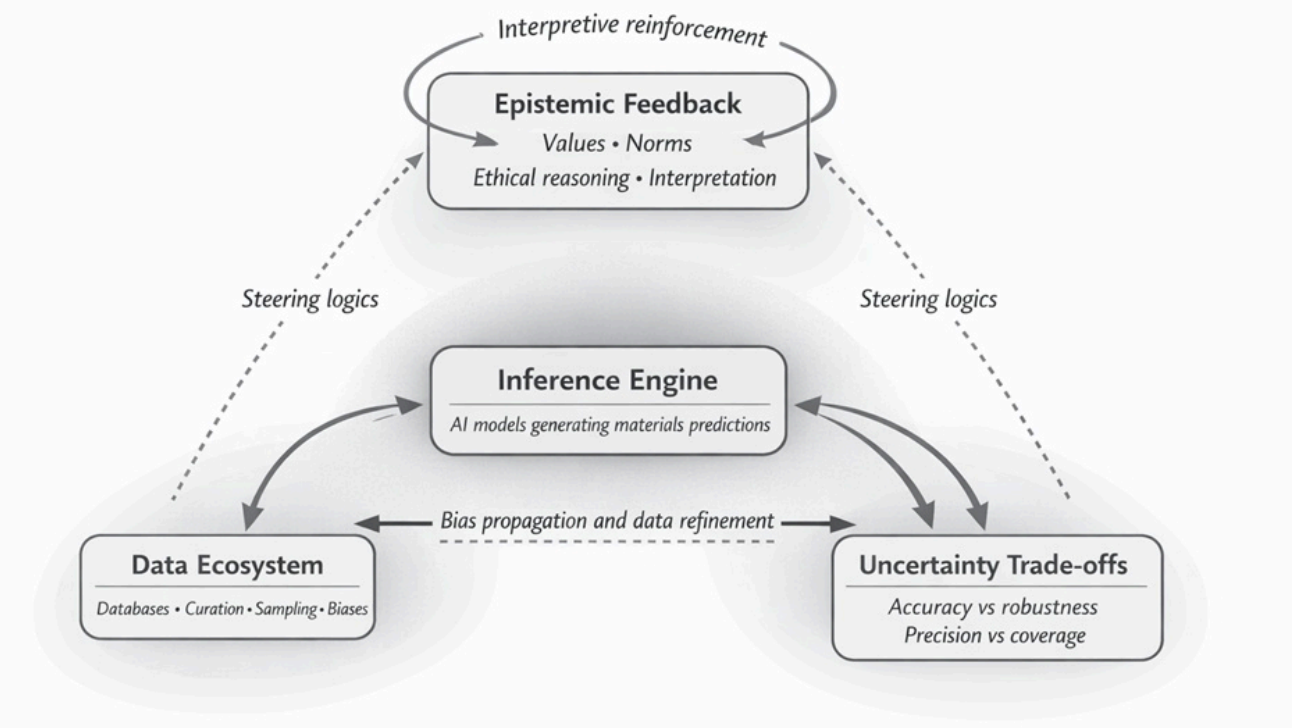

Figure 1. Conceptual framework of validation dynamics in AI-driven materials discovery

Figure 1 presents a schematic diagram of the core validation dynamics in AI-driven materials science. The framework centers on an Inference Engine (AI models generating predictions), which interacts dynamically with three interconnected conceptual domains: (1) the Data Ecosystem, representing databases and curation processes with inherent biases; (2) Epistemic Feedback, incorporating guiding values and ethical reasoning; and (3) Uncertainty Trade-offs, which manage analytical balances for reliability.

Bidirectional and looping arrows illustrate critical interactions: data biases influence inferences while model outputs inform data refinement; epistemic values create self-reinforcing cycles that steer other components; and explicit trade-offs modulate overall system behavior. Shaded gradients denote systems-level integration, with lighter zones indicating higher uncertainty. The triangular, integrated structure underscores the continuous, ground-truth-absent nature of validation in this domain.

The framework further explores conceptual interpretations of these dynamics, viewing validation as a process of epistemic harmonization. In practice, this involves navigating feedback structures to align inferences with scientific values, and addressing trade-offs between bias and uncertainty through reasoned integration [7, 28]. By doing so, it provides analytical tools for materials AI practitioners to enhance the conceptual integrity of inference.

Analytical implications

The conceptual framework developed in this study yields a set of analytical implications for understanding validation in materials AI under conditions where stable ground truth is unavailable or inaccessible. Rather than treating validation as a discrete methodological checkpoint, these implications foreground validation as an emergent systems property arising from interactions among inferential processes, epistemic values, and feedback structures. Collectively, they reframe validation as an interpretive practice shaped by steering logics, trade-offs, and recursive adjustment mechanisms.

Validation as an emergent, iterative process

A primary implication of the framework is that validation in materials AI cannot be analytically reduced to isolated model assessments. In the absence of empirical anchors, inferential credibility arises through iterative integrations of conceptual safeguards—such as data curation strategies, representational constraints, and uncertainty reasoning—distributed across the AI pipeline [1, 14]. Validation thus emerges over time, through repeated alignment between model behavior and epistemic expectations, rather than through a single benchmark comparison.

This perspective enables analytical interrogation of how bias propagation unfolds longitudinally. Decisions made at early stages—particularly regarding dataset scope and representativeness—act as steering logics that shape downstream inferential trajectories [22, 26]. As a result, validation becomes inseparable from the historical path-dependence of data and model co-evolution.

Data curation as a steering logic for inference

The framework highlights data curation not as a neutral preprocessing step but as a central epistemic control mechanism. When empirical ground truth is lacking, trade-offs between dataset comprehensiveness and representational balance directly influence the interpretive reliability of AI outputs [23, 27]. Analytical attention shifts from asking whether a model is "accurate" to examining *how* curated data configurations condition which inferences are possible or privileged.

This implication invites scholars to analyze validation failures not solely as algorithmic shortcomings but as outcomes of cumulative curatorial choices. Over time, such choices may stabilize narrow regions of materials space as epistemically "trusted," while marginalizing unexplored or underrepresented regimes.

Feedback structures and systems-level inferential integrity

At the systems level, the framework reveals feedback structures that regulate inferential integrity by linking model outputs, data augmentation strategies, and epistemic reassessment [6, 19]. Analytical exploration of these loops shows that validation operates through continuous recalibration rather than confirmation. Outputs inform future data-collection priorities, which, in turn, reshape the inferential landscape available to subsequent models.

This recursive dynamic underscores how ethical and epistemic reasoning become embedded within system behavior. For example, prioritizing algorithmic efficiency without parallel attention to representational diversity may intensify existing data biases, leading to progressively skewed conceptual interpretations of materials spaces [15, 21]. Validation, therefore, functions as a balancing act between exploratory freedom and constraint-driven caution.

Trade-Offs in uncertainty interpretation and risk management

Another analytical implication concerns the treatment of uncertainty. In ground-truth-deficient contexts, uncertainty is not merely a statistical descriptor but a central interpretive resource. The framework suggests that validation depends on how uncertainty is conceptually interpreted—whether it is used to temper confidence, guide exploration, or justify exclusion [7, 28].

This shifts analytical focus toward epistemic values that govern acceptable levels of risk. Overconfidence in unverified predictions can be understood as a systemic outcome of misaligned uncertainty interpretation rather than a localized modeling error. Validation, in this sense, reflects the degree to which epistemic humility is structurally embedded within AI-driven discovery processes.

Ethical–Epistemic integration in validation reasoning

The framework further implies that ethical considerations are analytically inseparable from epistemic validation. Choices about which materials to explore, which uncertainties to tolerate, and which inferences to prioritize are informed by value-

laden judgments embedded in system design and institutional context [30, 33]. Validation thus extends beyond technical coherence to include reflection on downstream implications and societal responsibility.

From this perspective, ethical reasoning functions as a stabilizing force within feedback structures, guiding harmonization across heterogeneous data sources and mitigating interpretive gaps produced by incomplete ground truth [32, 34]. Analytical engagement with validation, therefore, requires examining how values are operationalized within AI workflows rather than merely stated as external principles.

Validation as a networked interpretive practice

Finally, the framework positions validation as a network of interacting processes rather than a linear sequence. Inferential credibility is distributed across data representations, model assumptions, uncertainty interpretations, and value commitments [11, 25]. Analytical implications arise from mapping these interactions, revealing how coherence is maintained—or disrupted—across the system.

This networked view enables scholars to dissect validation failures as relational breakdowns rather than singular errors. It also provides conceptual tools for identifying leverage points where epistemic robustness can be strengthened without reliance on empirical benchmarking. These interactions are synthesized in **Table 1**, which conceptualizes validation in materials AI as an emergent systems-level property arising from coordinated trade-offs among data practices, inferential structures, uncertainty reasoning, and epistemic–ethical values.

Table 1. Integrative framework for validation as an emergent property in materials AI

Analytical dimension	Core conceptual mechanism	Role in validation dynamics	Dominant trade-offs	Epistemic risk if unmanaged
Data curation and representation	Dataset composition functions as a steering logic shaping accessible material space	Conditions that the model can plausibly infer in the absence of empirical ground truth	Coverage vs. representativeness; diversity vs. density	Reinforcement of historical bias; false confidence in narrow regimes
Model inference and representation learning	Inference emerges from learned latent structures rather than physical correspondence	Validation relies on internal coherence, stability, and cross-model agreement	Expressivity vs. interpretability; flexibility vs. constraint	Scientifically opaque predictions; spurious pattern stabilization
Ground truth substitution	Surrogate truth (simulations, consensus databases) replaces empirical anchors	Shifts validation from verification to alignment with accepted conventions	Speed vs. epistemic certainty	Propagation of systematic approximation errors
Uncertainty interpretation	Uncertainty operates as an interpretive signal rather than a corrective statistic	Modulates confidence, exploration, and exclusion decisions	Exploration vs. caution; decisiveness vs. humility	Overconfidence in unverified predictions; premature design lock-in
Feedback structures	Model outputs recursively influence data generation and prioritization	Validation becomes path-dependent and historically contingent	Efficiency vs. corrective capacity	Runaway feedback loops amplify early assumptions

Epistemic values	Reliability, transparency, and coherence guide acceptance of inferences	Substitute for benchmark validation in ambiguous regimes	Predictive power vs. explainability	Loss of scientific credibility
Ethical reasoning	Value judgments embedded in workflow design and prioritization	Aligns validation with societal and scientific responsibility	Innovation speed vs. risk containment	Skewed innovation trajectories; inequitable outcomes
Systems-level integration	Validation emerges from alignment across interacting components	Produces epistemic robustness without definitive ground truth	Local optimization vs. global coherence	Fragmented validation practices; inconsistent inferential standards

Results and Discussion

The interpretive lens provided by this conceptual framework enriches discussions on validation in materials AI by integrating analytical implications with broader epistemic and ethical reasoning. Central to this discourse is the recognition of interaction dynamics as foundational to inferential processes, where the lack of ground truth necessitates reliance on feedback structures for conceptual stability [2, 16]. This integration challenges conventional paradigms, suggesting that validation involves not merely technical adjustments but profound epistemic realignments, as biases in data ecosystems interact with scientific values to redefine reliability [24, 31].

Systems-level insights further illuminate the trade-offs inherent in AI applications for materials discovery, such as the tension between generative creativity and interpretive caution [5, 20]. Ethical reasoning emerges as a critical component, steering logics toward inclusive data practices that address historical skews, thereby fostering more equitable conceptual interpretations of chemical spaces [8, 17]. For instance, in high-throughput screening, these dynamics reveal how epistemic feedback can counteract bias amplification, promoting integrative approaches that enhance the overall epistemic fabric of materials science [10, 12].

Moreover, the framework's emphasis on conceptual interpretations invites reflection on the ethical dimensions of deployment, where steering logics must balance innovation with accountability [3, 27]. This discussion extends to the broader implications for scientific communities, highlighting how systems-level trade-offs influence collaborative knowledge production in AI-augmented environments [33, 35]. By interpreting validation through these multifaceted lenses, the framework contributes to ongoing dialogues on epistemic integrity, encouraging adaptive strategies that align AI inferences with enduring scientific values [28, 30].

In synthesizing these elements, the discussion underscores the transformative potential of conceptual frameworks in materials AI, where interaction dynamics and feedback structures offer interpretive pathways to overcome ground-truth limitations [21, 23]. This approach not only refines analytical tools but also enriches ethical epistemic reasoning, paving the way for more resilient inferential practices in the field [15, 18].

Conclusion

This manuscript has developed a novel conceptual theory for validation in materials AI, interpreting inference processes through interaction dynamics, systems-level insights, and epistemic feedback structures in the absence of ground truth. By synthesizing recent literature on materials informatics, data bias, and epistemic values, the framework offers analytical implications and interpretive tools for navigating trade-offs in uncertainty and bias. Ethical reasoning integrates throughout, steering logics toward enhanced conceptual coherence and inferential integrity.

Ultimately, this conceptual approach fosters deeper understandings of validation as an emergent, integrative process, contributing to applied AI in materials science by illuminating pathways for robust epistemic engagements amid data ambiguities.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 10 Apr 2021 Revised: 29 May 2021 Accepted: 05 Jul 2021
Published online: 18 January 2022

References

Madika B, Saha A, Kang C, Buyantogtokh B, Agar J, Wolverton CM, et al. Artificial intelligence for materials discovery, development, and optimization. ACS Nano. 2025;19(30):27116-58.
https://doi.org/10.1021/acsnano.5c04200.

Cheetham AK, Seshadri R. Artificial intelligence driving materials discovery? Perspective on the article: Scaling deep learning for materials discovery. Chem Mater. 2024.
https://doi.org/10.1021/acs.chemmater.4c00643.

Handoko AD, Made RI. Artificial intelligence and generative models for materials discovery: A review. arXiv. 2025;arXiv:2508.03278

Gregoire JM, Zhou L, Haber JA. Combinatorial synthesis for ai-driven materials discovery. Nat Synth. 2023;2:493-504.
https://doi.org/10.1038/s44160-023-00251-4.

Li J, Lim K, Yang H, Ren Z, Raghavan S, Chen PY. Ai applications through the whole life cycle of material discovery. Matter. 2020;3(2):393-432.
https://doi.org/10.1016/j.matt.2020.06.011.

Pyzer-Knapp EO, Pitera JW, Staar PWJ. Accelerating materials discovery using artificial intelligence, high performance computing and robotics. Npj Comput Mater. 2022;8(1):84.
https://doi.org/10.1038/s41524-022-00765-z.

Nematov D, Hojamberdiev M. Machine learning-driven materials discovery: Unlocking next-generation functional materials—a minireview. SSRN. 2025.
https://doi.org/10.2139/ssrn.5219988.

Fuhr AS, Sumpter BG. Deep generative models for materials discovery and machine learning-accelerated innovation. Front Mater. 2022;9:865270.
https://doi.org/10.3389/fmats.2022.865270.

Li C, Zheng K. Methods, progresses, and opportunities of materials informatics. InfoMat. 2023;5(6):e12425.
https://doi.org/10.1002/inf2.12425.

Hart M, Idanwekhai K, Alves VM, Miller AJM. Trust not verify? The critical need for data curation standards in materials informatics. Chem Mater. 2024.
https://doi.org/10.1021/acs.chemmater.4c00981.

Frydrych K, Karimi K, Pecelerowicz M, Alvarez R. Materials informatics for mechanical deformation: A review of applications and challenges. Materials. 2021;14(19):5764.
https://doi.org/10.3390/ma14195764.

Moran M. An information oriented approach to materials informatics. Liverpool: University of Liverpool Repository; 2025.

Agrawal A, Gopalakrishnan K. Materials image informatics using deep learning. Singapore: World Scientific; 2020.

Kumagai M, Ando Y, Tanaka A, Tsuda K. Effects of data bias on machine-learning–based material discovery using experimental property data. Adv Mater Interfaces. 2022.
https://doi.org/10.1002/admi.202200944.

Ntoutsi E, Fafalios P, Gadiraju U. Bias in data-driven artificial intelligence systems—an introductory survey. WIREs Data Min Knowl Discov. 2020;10(3):e1356.
https://doi.org/10.1002/widm.1356.

Trezza G, Chiavazzo E. Classification-based detection and quantification of cross-domain data bias in materials discovery. J Chem Inf Model. 2024;65(4):1747-61.
https://doi.org/10.1021/acs.jcim.4c00012.

Leavy S, O'Sullivan B, Siapera E. Data, power and bias in artificial intelligence. arXiv. 2020;arXiv:2008.07341.

Zhang H, Chen W, Rondinelli J. Mitigating bias in scientific data: A materials science case study. AI for Science Workshop. 2023.

Davariashtiyani A, Wang B, Hajinazar S. Impact of data bias on machine learning for crystal compound synthesizability predictions. Mach Learn Sci Technol. 2024;5(4):045001.
https://doi.org/10.1088/2632-2153/ad9378.

Tejani AS, Ng YS, Xi Y, Rayan JC. Understanding and mitigating bias in imaging artificial intelligence. Radiographics. 2024;44(1):e230067.
https://doi.org/10.1148/rg.230067.

Daneshjou R, Smith MP, Sun MD. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: A scoping review. JAMA Dermatol. 2021;157(11):1362-9.
https://doi.org/10.1001/jamadermatol.2021.3129.

Zhang Z, Li J, Stork DG, Mansfield E. Reducing bias in ai-based analysis of visual artworks. IEEE BITS the International Week. 2022.

Dehkharghanian T, Bidgoli AA, Riasatian A. Biased data, biased ai: Deep networks predict the acquisition site of tcga images. Diagn Pathol. 2023;18(1):86.
https://doi.org/10.1186/s13000-023-01355-3.

Alvarado R. Ai as an epistemic technology. Sci Eng Ethics. 2023;29(5):1-30.
https://doi.org/10.1007/s11948-023-00451-3.

Cheung KKC, Long Y, Liu Q, Chan HY. Unpacking epistemic insights of artificial intelligence (ai) in science education: A systematic review. Sci Educ. 2024;34(2):747-77.
https://doi.org/10.1007/s11191-024-00511-5.

Cruz-Aguilar MA. The epistemic revolution of ai: Reconfiguring the foundations of scientific knowledge. AI Soc. 2025.
https://doi.org/10.1007/s00146-025-02658-3.

Russo F, Schliesser E, Wagemans J. Connecting ethics and epistemology of ai. AI Soc. 2024;39(2):361-3.
https://doi.org/10.1007/s00146-022-01617-6.

Ngwenyama O, Rowe F. Should we collaborate with ai to conduct literature reviews? Changing epistemic values in a flattening world. J Assoc Inf Sci Technol. 2024;75(1):67-82.
https://doi.org/10.1002/asi.24835.

Berber A, Mijić J. Epistemic education for an ai-driven world. EMERGE 2024 Ethics of AI Alignment—Book of Abstracts. 2024

Koskinen I. We have no satisfactory social epistemology of ai-based science. Soc Epistemol. 2024;38(4):458-75.
https://doi.org/10.1080/02691728.2023.2286253.

Hauswald R. Artificial epistemic authorities. Soc Epistemol. 2025;39(6):1-10.
https://doi.org/10.1080/02691728.2025.2449602.

Lei K, Joress H, Persson N. Aggressively optimizing validation statistics can degrade interpretability of data-driven materials models. J Chem Phys. 2021;155(5):054105.
https://doi.org/10.1063/5.0052205.

Hartung T, Kleinstreuer N. Challenges and opportunities for validation of ai-based new approach methods. ALTEX. 2025

Myllyaho L, Raatikainen M, Männistö T. Systematic literature review of validation methods for ai systems. arXiv. 2021;arXiv:2107.12190

Ebrahimian S, Kalra MK, Agarwal S, Bizzo BC. Fda-regulated ai algorithms: Trends, strengths, and gaps of validation studies. Acad Radiol. 2022;29(6):805-11.
https://doi.org/10.1016/j.acra.2021.09.018.