ORIGINAL RESEARCH **Open access**

Error Propagation Across the Materials AI Pipeline: A System-Level Conceptual Model

Neha Sharma[1], Ravi Kumar[1*], Aniket Deshmukh[2], Arjun Nair[1]

Abstract

The integration of artificial intelligence (AI) into materials science has transformed the landscape of material discovery and design, enabling accelerated exploration of vast compositional spaces and property predictions. However, this advancement introduces complex error dynamics that permeate the entire AI pipeline, from data curation to model deployment. This conceptual manuscript develops a novel system-level model to interpret error propagation within materials AI workflows, emphasizing interaction dynamics and feedback structures rather than empirical validation. Drawing on recent literature, the synthesis reveals fragmented understandings of error sources, such as data inconsistencies and algorithmic biases, and their cascading effects across pipeline stages. The proposed framework conceptualizes the pipeline as an interconnected system where errors manifest through amplification, mitigation, and transformation mechanisms, informed by epistemic considerations and trade-off analyses. Analytical implications highlight how these dynamics influence interpretability and reliability in materials innovation, while ethical reasoning underscores the need for holistic oversight. By integrating conceptual interpretations from uncertainty quantification and systems theory, this model offers insights into steering logics that balance precision with robustness, fostering a deeper understanding of AI's role in advancing materials science without relying on testable claims or experimental data.

Keywords Artificial intelligence, Materials science, Conceptual frameworks, Uncertainty dynamics, Error propagation, System interactions

*Correspondence:
Ravi Kumar
ravi.kumar@gmail.com
[1] Department of Materials Engineering and Data Science, Faculty of Engineering, IIT Delhi, New Delhi, India
[2] Department of AI-Based Materials Design, Faculty of Engineering, IIT Bombay, Mumbai, India

Introduction

The emergence of artificial intelligence (AI) within materials science marks a profound epistemic and methodological transformation in how materials are designed, discovered, and understood. Traditionally, progress in materials science has relied on iterative experimentation, physically grounded simulations, and expert intuition—processes that, while powerful, are often constrained by high costs, long development cycles, and the combinatorial complexity of composition–structure–property spaces. Over the past decade, advances in AI, particularly machine learning (ML), have been positioned as a response to these limitations, offering data-driven mechanisms for navigating vast materials design spaces with unprecedented speed and apparent predictive accuracy [1, 2].

In contemporary materials research, AI is rarely deployed as a single algorithmic artifact; rather, it operates through multi-stage pipelines that integrate heterogeneous data sources, representation choices, learning architectures, and decision heuristics. Typical materials AI workflows encompass data acquisition from experimental repositories or computational

simulations, feature engineering or representation learning, model training and validation, inference, and iterative refinement through active learning or feedback from domain experts [3]. Within these pipelines, AI models are increasingly tasked with predicting diverse material attributes—ranging from mechanical strength and thermal stability to electronic bandgaps and catalytic activity—often under conditions where direct experimentation would be infeasible. Conceptual advances such as graph neural networks for atomistic representation and generative models for chemical space exploration exemplify how AI is reshaping not only computational efficiency, but also the representational language of materials science itself [4, 5].

Despite these advances, the growing reliance on AI-generated insights has brought renewed attention to a foundational concern: the reliability and epistemic status of AI-mediated knowledge claims in materials science. Crucially, reliability challenges do not arise solely from isolated coding errors or numerical instabilities, but from systematic deviations that emerge across the entire pipeline. In this context, error refers to any structured distortion—conceptual, representational, or inferential—that alters the fidelity between AI outputs and the material phenomena they purport to describe [6]. Such errors may originate in data incompleteness, biased sampling, model misspecification, or mismatches between physical reality and learned abstractions, and they frequently interact in non-trivial ways.

Recent scholarly discourse has increasingly acknowledged that these errors do not remain localized but instead propagate through interconnected stages of materials AI workflows [7]. For example, training datasets assembled from heterogeneous sources—such as density functional theory (DFT) calculations, high-throughput simulations, and experimental measurements—often embed implicit assumptions, approximations, and noise profiles that are not uniformly documented or harmonized [8]. When such datasets are ingested into learning pipelines without explicit epistemic accounting, biases introduced at the data level may cascade into model behavior, downstream predictions, and ultimately into scientific interpretations of material behavior.

Understanding this phenomenon requires a shift from component-level diagnostics toward a system-level perspective on materials AI. Drawing on systems theory, this view emphasizes that complex pipelines exhibit emergent behavior arising from interactions among subsystems rather than from any single element in isolation [9]. Within materials AI, these interactions frequently manifest as feedback loops: preprocessing choices influence representational learning; representational constraints shape model inductive biases; and model outputs, in turn, guide subsequent data acquisition or experimental prioritization [10]. Errors introduced at early stages may thus be amplified, transformed, or obscured as they traverse the pipeline, complicating efforts to trace their origins or assess their impact.

Importantly, this manuscript does not seek to quantify error propagation empirically, nor to propose new statistical metrics. Instead, it advances a qualitative, conceptual analysis of the structural mechanisms through which errors flow, interact, and acquire epistemic significance within materials AI systems [11]. The central challenge addressed here is interpretive: how to reason about error dynamics when AI pipelines function as coupled socio-technical systems that blend data, algorithms, physical assumptions, and human judgment.

The need for such a conceptual model is underscored by the fragmented treatment of error in the existing literature. Many studies examine uncertainty, bias, or robustness in isolation—focusing, for instance, on algorithmic uncertainty quantification or dataset curation practices—without integrating these insights into a coherent pipeline-level account [12]. While notable progress has been made in uncertainty quantification for ML models [13], these approaches are rarely contextualized within the distinctive constraints of materials science, where multiscale phenomena, physical laws, and limited experimental access introduce additional layers of complexity [14]. As a result, expectations regarding the reliability and transferability of AI-assisted materials insights are often misaligned with the structural realities of the pipelines that produce them [15].

By conceptualizing error propagation as a systemic property rather than a localized defect, this work aims to clarify the trade-offs inherent in materials AI design choices. These include tensions between model complexity and interpretability, predictive accuracy and robustness, and data efficiency and epistemic coverage. Explicitly framing these trade-offs

enables more reflective deployment of AI tools and supports stronger forms of scientific reasoning about when and how AI-derived insights should be trusted [16].

Beyond epistemic considerations, error propagation also carries ethical and societal implications. Pipelines trained predominantly on well-curated, resource-intensive datasets risk reinforcing inequities by marginalizing underrepresented material classes, regions, or sustainability-oriented research agendas [17]. From a conceptual standpoint, ethical governance of materials AI therefore requires steering logics that foreground transparency, accountability, and inclusivity across the entire pipeline—not merely at the point of model deployment [18]. This concern is particularly salient in emerging domains such as sustainable and green materials design, where AI-guided decisions may influence long-term environmental outcomes and resource allocation [19].

The remainder of this manuscript is structured as follows. The next section synthesizes recent literature to establish a theoretical background on AI pipelines in materials science, major sources of error, and prevailing conceptual assumptions. This is followed by the presentation of a novel system-level conceptual framework that explicates interaction dynamics and feedback structures governing error propagation. The discussion section elaborates the epistemic, methodological, and ethical implications of this framework, and the concluding section summarizes key insights and outlines future directions for conceptual and governance-oriented research in materials AI.

In summary, this work offers a purely conceptual contribution grounded in critical synthesis rather than empirical evaluation. By articulating error propagation as an emergent, system-level phenomenon, it seeks to enrich the interpretive vocabulary available to researchers and practitioners deploying AI in materials science. Through this lens, AI is not merely a tool for acceleration, but a transformative epistemic instrument whose reliability depends on how its internal structures, assumptions, and interactions are understood and governed [20].

Theoretical Background and Literature Synthesis

AI applications in materials science

Since 2020, the application of artificial intelligence in materials science has accelerated markedly, motivated by the need to shorten discovery cycles in domains such as energy storage, catalysis, functional polymers, and biomaterials [21]. Rather than functioning as a mere optimization aid, AI is increasingly conceptualized as an epistemic mediator—a system that reshapes how material hypotheses are generated, evaluated, and prioritized. Contemporary frameworks frequently cast AI as an enabler of inverse design, wherein desired material properties are specified first, and algorithmic models explore candidate compositions or structures that satisfy these constraints [22].

Within this paradigm, deep learning architectures have gained prominence for their ability to operate on high-dimensional, heterogeneous data. Conceptually, these models are understood as mapping complex structure–property relationships that would be intractable under classical analytical approaches, particularly in systems such as high-entropy alloys, polymer networks, and multicomponent oxides [23]. Importantly, the literature emphasizes that these successes are not attributable to isolated models but to integrated pipelines that connect atomistic simulations, curated databases, representation learning, and predictive inference. Such pipelines are often framed as hybrid constructs that fuse physics-based reasoning with data-driven abstraction, creating a productive—yet tension-laden—interface between theoretical materials science and machine learning [24].

At the same time, conceptual analyses caution that AI performance is inseparable from the fidelity of material representations. Studies synthesizing work from 2021–2023 show that graph-based and message-passing models encode atomic environments in ways that are highly sensitive to representation choices, hyperparameters, and dataset composition [25]. While these models are celebrated for their ability to bridge length scales—from atomic bonding to bulk properties—systems-level perspectives stress that such bridging is interpretive rather than neutral. Misalignments

between representational abstractions and physical reality can subtly distort scientific understanding, even when predictive accuracy appears high [26].

Sources of error in AI models for materials

Across the literature, error in materials AI is no longer treated as a residual nuisance but as a structural feature of learning systems operating under epistemic constraints. Conceptual syntheses distinguish between epistemic uncertainty, arising from incomplete knowledge or limited coverage of materials space, and aleatoric uncertainty, linked to intrinsic variability in measurements or stochastic processes [27]. Both forms are deeply entangled in materials AI workflows and shape not only numerical outputs but also their interpretive meaning.

Data-related errors are consistently identified as the dominant source of downstream distortion. Curated materials databases frequently aggregate results from density functional theory calculations, high-throughput simulations, and experimental measurements—each governed by distinct assumptions, approximations, and noise profiles [28]. Conceptual analyses published since 2022 highlight how these heterogeneities give rise to latent biases that are often invisible during model training yet materially consequential for inference [29]. In particular, descriptor incompleteness and representational compression may erase quantum-mechanical or microstructural nuances, producing predictions that are internally consistent but physically misleading.

Algorithmic sources of error further compound these challenges. Overfitting, architectural inductive bias, and approximation error in deep neural networks have been reframed in the literature as semantic transformations, whereby the meaning of a predicted property shifts relative to its physical counterpart [30]. Rather than viewing robustness solely as a statistical attribute, recent conceptual work interprets robustness as an epistemic condition—one that governs whether AI outputs can legitimately support material reasoning and design decisions [31]. Ethical discussions intersect with this perspective by emphasizing trade-offs between computational efficiency, model opacity, and the responsibility to minimize systematic distortion, particularly when AI systems influence high-stakes material choices [32].

Pipeline structures in materials AI workflows

The prevailing theoretical view of materials AI pipelines conceptualizes them as sequential yet tightly coupled systems that span data ingestion and preprocessing through modeling, validation, and deployment [31]. Literature from 2020 onward increasingly rejects linear depictions in favor of dynamic models that emphasize feedback, iteration, and cross-stage dependency [32]. Within such systems, decisions made at early stages—such as data filtering, normalization, or dimensionality reduction—are understood to exert disproportionate influence on downstream representations and predictions.

Conceptual analyses underscore that preprocessing is not a neutral technical step but an epistemic intervention. Dimensionality reduction and feature selection, while essential for tractability, may suppress physically meaningful variance, thereby reshaping the hypothesis space accessible to learning algorithms [33-35]. More recent frameworks describe advanced pipelines as adaptive networks that blend supervised, unsupervised, and generative components, allowing models to respond dynamically to domain-specific constraints and evolving data regimes [31].

At the system level, interaction dynamics between pipeline stages play a decisive role in error propagation. Validation and ensemble strategies are often interpreted as corrective mechanisms that dampen upstream errors; however, literature cautions that these same mechanisms can also mask structural deficiencies, enabling confident deployment of fragile models [35]. As a result, contemporary theoretical syntheses argue for viewing materials AI pipelines as epistemic infrastructures—systems whose reliability depends not only on individual components, but on how assumptions, uncertainties, and corrective mechanisms interact across the entire workflow.

Uncertainty quantification and error dynamics

Uncertainty quantification (UQ) has emerged as a central conceptual instrument for interpreting error behavior in materials AI, moving the discourse beyond point predictions toward epistemically qualified claims [33]. Rather than serving merely as a statistical add-on, UQ is increasingly framed as a representational layer that mediates how confidence, reliability, and limitation are communicated within AI-driven materials pipelines. Contemporary syntheses emphasize probabilistic approaches—such as Bayesian inference, ensemble learning, and posterior predictive distributions—not primarily for numerical refinement, but for their capacity to make uncertainty structurally visible within predictive workflows [34].

From a systems perspective, uncertainty is not static or localized. Recent conceptual work highlights how uncertainties introduced at specific stages—such as noisy experimental measurements or approximated simulation outputs—can interact and transform as they propagate through learning architectures [35]. Local perturbations in input space may thus manifest as global distortions in downstream inference, particularly when models are extrapolated beyond their effective domains of applicability. This transformation challenges simplistic interpretations of uncertainty as additive error and instead positions it as a dynamic property of pipeline interactions.

Conceptual interpretations further stress the role of feedback structures in moderating uncertainty dynamics. In iterative materials AI workflows, uncertainty estimates increasingly inform decisions about data acquisition, model refinement, and experimental prioritization. These feedback loops are framed as balancing mechanisms that trade computational efficiency against epistemic caution, shaping how aggressively models are trusted or revised [32]. Importantly, such balances are not value-neutral: they encode implicit judgments about acceptable risk, confidence thresholds, and the relative cost of error in materials discovery contexts.

Ethical reasoning is therefore inseparable from UQ in materials AI. Recent theoretical discussions argue that acknowledging uncertainty is not merely a technical obligation but an epistemic responsibility, particularly when AI outputs guide high-stakes scientific or sustainability-related decisions [32]. Concealing or under-communicating uncertainty may artificially inflate confidence, while overly conservative representations may impede innovation. Conceptual clarity around uncertainty thus becomes a prerequisite for responsible scientific reasoning rather than a secondary reporting concern.

Integration of systems theory in error analysis

Systems theory offers a powerful conceptual framework for synthesizing error propagation in materials AI as an emergent phenomenon arising from interactions among pipeline components rather than from isolated failures [31]. Since 2021, the literature increasingly conceptualized materials AI pipelines as complex adaptive systems, characterized by nonlinear dynamics, feedback loops, and path-dependent behavior. Within this framing, error flows are governed not only by the magnitude of local inaccuracies but by how structural couplings amplify, dampen, or redirect them across stages.

A central insight from this synthesis is the recognition of trade-offs between modularity and holistic coherence. Highly modular pipelines may simplify debugging and optimization at individual stages, yet risk obscuring cross-stage dependencies that give rise to emergent error patterns. Conversely, tightly integrated systems may achieve greater conceptual alignment between data, models, and physical assumptions, but at the cost of reduced transparency and interpretability. Systems-theoretic analyses interpret these trade-offs as steering logics—design choices that shape the overall reliability, adaptability, and epistemic robustness of materials AI workflows [32]. To consolidate the literature-derived error sources, propagation mechanisms, and epistemic consequences across pipeline stages, **Table 1** provides a system-level synthesis mapping how errors originate, transform, and are governed throughout materials AI workflows.

Table 1. System-level synthesis of error propagation across the materials AI pipeline: sources, transformation mechanisms, epistemic consequences, and steering logics

Pipeline stage	Typical inputs and operations	Primary error/uncertainty	Propagation and	Observable symptoms	Miti res

		sources	transformation mechanisms	(diagnostic signals)	mechar + cro
1) Data curation	Experimental repositories; DFT/high-throughput simulations; literature-extracted data; metadata harmonization; unit standardization	Heterogeneous provenance; untracked simulation settings; measurement noise; missing regimes; selection bias; class imbalance; label uncertainty	Foundation bias: distorted priors enter all downstream stages; dataset shift across sources; silent mismatch between “same” property definitions across labs/simulations; error seeding (early small deviations become structural)	Inconsistent distributions across sources; high variance across nominally similar samples; sensitivity to train/test split; model performance collapses under external validation	Provena standar samp fidelit outlie dataset definiti property applic
2) Feature engineering/representations	Hand-crafted descriptors; graph encodings; structural fingerprints; dimensionality reduction; normalization; feature selection	Descriptor incompleteness; physically invalid invariances; correlated features; representation collapse; leakage (target information encoded unintentionally)	Semantic distortion: representation changes what the model can “mean”; amplification via correlation (spurious proxies become dominant); compression loss removes physically relevant variance	High performance with low physical plausibility; unstable feature importance; contradictory explanations across similar samples; sensitivity to normalization choices	Physic feature checks; feature t s represer uncert eml
3) Model construction and training	Supervised ML; GNNs; ensembles; generative models; hyperparameter search; regularization; validation protocols	Overfitting; inductive bias misalignment; optimization instability; miscalibration; objective mismatch (loss ≠ scientific goal)	Error internalization: models learn and reinforce upstream bias; overconfidence dynamics through miscalibration; robustness illusion when validation is not	Poor calibration; high confidence on out-of-domain points; performance depends strongly on random seed; high variance across folds;	Calibrat Bayesia robus (OC constra uncert objectiv analy object

			distribution-aware	fragile generalization	
4) Inference, explanation, and interpretation	Predictions for new compositions/structures; explanation methods; surrogate reasoning; expert interpretation; decision thresholds	Extrapolation beyond applicability; explanation instability; misinterpretation of uncertainty; post-hoc rationalization	Narrative coupling: outputs become scientific "stories"; explanation drift (explanations change with minor perturbations); threshold effects convert uncertainty into hard decisions	Contradictory explanations for similar inputs; brittle rankings; high sensitivity to small input changes; divergence between expert intuition and model rationale	Do applica uncert abstenti conse explan (stabil human revie
5) Iterative deployment and feedback	Active learning; closed-loop experimentation; retraining with new data; operational constraints; continual updates	Feedback bias (only "successful" results recorded); automation bias; concept drift; reward hacking (optimize proxy outcomes)	Reinforcement of blind spots: system learns its own sampling bias; nonlinear escalation where small early biases reshape future data; lock-in to narrow regions of design space	Narrowing diversity of suggested candidates; declining novelty; escalating confidence without improved external validity; drift in model behavior over time	Exp constrai regulari externa cha re goverr trails; rc

From an epistemic standpoint, this literature reframes error not solely as a defect to be eliminated but as an informative signal that reveals structural tensions within the system. Patterns of recurring failure, instability, or uncertainty concentration can illuminate mismatches between representation, learning objectives, and physical reality. Conceptual models integrating systems theory with materials-specific challenges—such as multiscale coupling and limited experimental access—thus provide guidance for designing pipelines that are not error-free, but error-aware and resilient.

Proposed conceptual framework

The proposed conceptual framework interprets error propagation in materials AI pipelines through a system-level lens, emphasizing interaction dynamics, feedback structures, and epistemic trade-offs. This model conceptualizes the pipeline as a multilayered network comprising five core stages: data curation, feature engineering, model construction, inference and interpretation, and iterative deployment. Errors are viewed as fluid entities that traverse these stages, undergoing transformations influenced by inter-stage dependencies and systemic feedback.

At the data curation stage, errors originate from heterogeneous sources, such as discrepancies between simulated and experimental datasets, manifesting as inconsistencies that set the foundational tone for subsequent dynamics. These initial errors interact with feature engineering, where selection and transformation processes can either amplify distortions —by inadequately representing physical constraints—or mitigate them through domain-informed filtering. The framework

highlights trade-offs: pursuing comprehensive feature sets may introduce computational complexity that, in turn, indirectly fosters error accumulation.

Transitioning to model construction, errors from prior stages are integrated into algorithmic architectures, influencing learning paradigms such as supervised or generative models. Conceptual interpretations reveal how biases in optimization routines transform localized errors into systemic variances, with feedback loops from validation metrics steering adjustments that balance accuracy against generalizability. Epistemic reasoning underscores the interpretive challenge: errors at this juncture can skew the conceptual mapping of material phenomena, potentially distorting understandings of structure-property linkages.

In the inference and interpretation stage, propagated errors manifest in output uncertainties, where systems-level insights illuminate their role in shaping scientific narratives. Interaction dynamics between predicted properties and domain knowledge create opportunities for error attenuation through ensemble approaches, yet also risk escalation if unaddressed. The framework interprets these as ethical imperatives, prompting steering logics that prioritize transparency in conveying uncertainty to end-users.

Finally, iterative deployment encapsulates the pipeline's cyclical nature, with errors feeding back into data curation via real-world applications, fostering adaptive evolutions. Analytical implications suggest that such feedback establishes resilience thresholds, in which trade-offs between innovation speed and error control define the long-term epistemic value. The integrative structure and error flow across the materials AI pipeline are visualized in **Figure 1**.

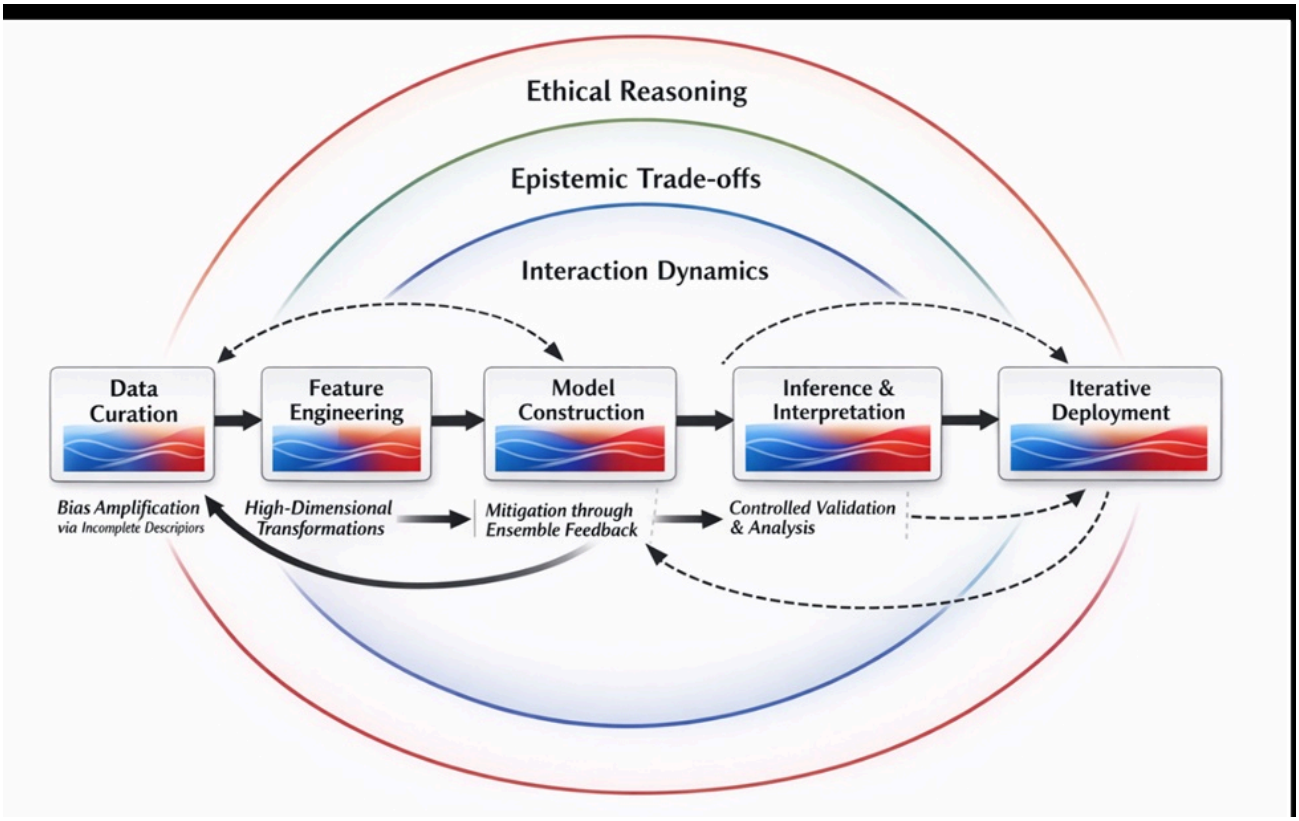

Figure 1. Schematic representation of the materials AI pipeline emphasizing interconnected stages and error propagation dynamics.

Analytical implications

The conceptual framework advanced in this manuscript provides a systematic lens for interpreting the complex and often opaque dynamics of error propagation in materials AI pipelines. Rather than treating errors as isolated artifacts attributable to individual stages, the framework foregrounds interaction dynamics—how errors are transformed, redistributed, and reinterpreted as they traverse interconnected pipeline components. For example, deviations arising from data curation may undergo qualitative transformations during feature engineering, where domain-specific preprocessing, normalization, or descriptor construction can either dampen or amplify initial distortions. Mechanisms such as feature correlation dependencies or representational compression play a decisive role in this transformation, reshaping how material information is encoded and subsequently learned by models [6].

Crucially, these transformations are not linear or unidirectional. Feedback structures linking inference outcomes to upstream stages introduce adaptive pathways that continuously recalibrate data selection, feature construction, and model refinement. Such feedback loops allow pipelines to balance competing objectives—error mitigation, computational efficiency, and predictive coverage—by redistributing uncertainty rather than eliminating it outright [10]. From a systems-level perspective, this redistribution gives rise to emergent properties, including enhanced robustness in hybrid or ensemble pipelines, where uncertainty is diffused across layers instead of accumulating at a single point of failure [9]. These properties cannot be inferred from individual components alone, underscoring the necessity of holistic analytical reasoning.

The framework further elucidates the epistemic trade-offs embedded in pipeline design choices. High-fidelity models, particularly those employing deep or highly expressive architectures, are often pursued to improve interpretability or capture subtle structure–property relationships. Paradoxically, such models may increase vulnerability to biases propagated from earlier stages, as their expressive capacity can internalize and reinforce latent distortions in the training data [16]. Steering logics, therefore, become essential: they guide decisions regarding when to privilege precision over generalizability, or interpretability over coverage. This is especially salient in materials discovery contexts where epistemic uncertainty is high—such as predicting properties for chemically novel or sparsely sampled compositions—demanding conservative modeling strategies to avoid unwarranted confidence [13].

Ethical reasoning deepens these analytical implications by situating error propagation within broader research equity concerns. Unchecked amplification of error can systematically disadvantage underrepresented material classes or research domains, including rare-earth alternatives and low-data sustainability materials, where epistemic gaps are already pronounced [17]. From this standpoint, error dynamics are not merely technical challenges but structural factors that can reinforce or mitigate disparities in scientific attention and resource allocation. The framework, therefore, promotes integrative thinking, encouraging researchers to treat trade-offs as opportunities for holistic optimization rather than zero-sum compromises, and to design pipelines that align technical performance with ethical imperatives of transparency, inclusivity, and accountability [18].

The emphasis on feedback structures also yields important analytical insights into pipeline resilience. In iterative deployment scenarios, errors evolve through cyclic interactions between prediction, validation, and refinement stages. Rather than being static defects, errors become dynamic signals that can trigger self-correcting behaviors over time, contributing to improved long-term reliability [34]. Interpreted through a systems-theoretic lens, such behaviors exemplify adaptive learning, where downstream performance metrics inform upstream adjustments—such as recalibrating feature sets, revising representation choices, or redefining training objectives [29]. This view reframes error as a diagnostic resource, revealing structural tensions and misalignments that would remain hidden under purely performance-oriented evaluation.

At a higher level of abstraction, the framework highlights the nonlinear nature of error propagation in materials AI systems. Small perturbations introduced during data curation—such as biased sampling or inconsistent simulation settings—can induce disproportionate shifts in inference outcomes once propagated through tightly coupled pipeline stages. These nonlinear effects prompt epistemic reflection on the limits of AI in capturing multiscale materials phenomena, particularly when models are asked to bridge quantum-mechanical approximations and macroscopic

property predictions [14]. Design trade-offs between modularity and integration become especially salient in this context: modular pipelines facilitate localized error isolation and interpretability, whereas tightly integrated pipelines promote coherence and feedback at the expense of increased complexity and reduced transparency [26].

Finally, the analytical implications extend beyond technical design to materials innovation and governance contexts. In sustainable materials design, for instance, the framework clarifies how error dynamics shape the credibility of AI-driven assessments of environmental impact or lifecycle performance. This underscores the need for steering logics that explicitly incorporate uncertainty transparency, enabling policymakers and stakeholders to make informed decisions grounded in epistemically qualified predictions rather than overconfident projections [19]. More broadly, the analysis reveals that error transformation mechanisms can function as strategic indicators, identifying domains—such as quantum-mechanical modeling or poorly constrained experimental regimes—where epistemic uncertainty dominates and targeted investment in data generation or model refinement is most warranted [28].

Taken together, these analytical implications reinforce a view of AI as a collaborative epistemic partner in materials science rather than an autonomous decision-maker. Human oversight, domain expertise, and ethical judgment remain indispensable for interpreting AI outputs, navigating trade-offs, and steering pipelines toward scientifically and socially responsible outcomes. By foregrounding error dynamics as a central analytical concern, the framework supports a more reflective, resilient, and trustworthy integration of AI into materials research practice [11].

Results and Discussion

The system-level conceptual model advanced in this manuscript integrates fragmented insights from recent literature on AI in materials science, offering a cohesive interpretive framework for error propagation that transcends individual component analyses [1]. By focusing on interaction dynamics and feedback structures, the model complements existing conceptualizations of uncertainty quantification, which often emphasize algorithmic techniques without fully addressing their pipeline-wide ramifications [31]. For example, while Bayesian methods have been synthesized as tools for handling epistemic uncertainties [18], the current framework interprets their role within feedback loops, where they facilitate adaptive transformations of errors across stages, enhancing systems-level resilience [12]. This integrative approach reveals trade-offs in adopting such methods, as increased computational demands may conflict with the need for real-time deployment in materials discovery workflows [4].

Conceptual interpretations of the model highlight its potential to bridge gaps in the literature, particularly regarding how errors interact with materials-specific constraints, such as multiscale modeling [14]. Literature syntheses on graph neural networks for atomic representations [25] can be reinterpreted through this lens, showing how biases in structural descriptors propagate and interact with model training dynamics, leading to transformed uncertainties in property predictions [11]. Ethical reasoning is integral to this discussion, as the framework underscores the risks of error amplification in biased datasets, which could perpetuate inequities in materials research priorities [17]. Steering logics, therefore, become essential for balancing innovation with accountability, encouraging pipeline designs that incorporate diversity in data sources to mitigate systemic distortions [12].

Systems-level insights from the model also inform discussions on robustness in emerging AI applications, such as generative models for chemical space exploration [5]. The framework's emphasis on error mitigation through ensemble feedback structures aligns with recent conceptual advances in hybrid learning paradigms [16]. Still, it extends them by exploring how these structures interact with epistemic trade-offs, such as those between model expressivity and error control [30]. This integrative reasoning suggests that error propagation is not merely a technical challenge but an opportunity for epistemic advancement, where understanding transformation mechanisms can refine conceptual mappings of material behaviors [27]. Furthermore, the model's focus on cyclical deployment feedback resonates with applications of systems theory in AI pipelines [13], offering interpretive depth into how nonlinear dynamics govern error flows in complex materials systems [14].

The framework's originality lies in its purely conceptual nature, avoiding empirical validation and instead providing interpretive tools for navigating pipeline complexities [7]. This approach addresses gaps in the literature on holistic error analysis, where siloed approaches to data quality or algorithmic bias fail to capture cascading effects [12]. By emphasizing ethical and epistemic dimensions, the model promotes a nuanced view of AI's role in materials science, framing trade-offs as opportunities for sustainable progress [15]. Future conceptual extensions could explore integration with quantum computing paradigms, interpreting how quantum uncertainties intersect with classical error dynamics in hybrid pipelines [4]. Overall, this discussion underscores the model's contribution to fostering integrative, systems-oriented thinking in applied AI for materials innovation [20].

Conclusion

This conceptual manuscript has developed a system-level model for interpreting error propagation in materials AI pipelines, drawing on interaction dynamics, feedback structures, and epistemic trade-offs to provide a holistic lens. By synthesizing recent literature and articulating the transformative nature of errors across stages, the framework offers interpretive insights that enhance understanding of AI's reliability in materials science. Ethical reasoning and steering logics emerge as key elements, guiding balanced approaches to innovation amid inherent uncertainties. Ultimately, this work contributes to a more integrated appreciation of AI workflows, paving the way for conceptually robust advancements in the field.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 02 Jul 2024 Revised: 12 Aug 2024 Accepted: 26 Sep 2024
Published online: 18 January 2025

References

Nemani V, Biggio L, Huan X, Hu Z, Fink O, Tran A, et al. Uncertainty quantification in machine learning for engineering design and health prognostics: A tutorial. Mech Syst Signal Process. 2023;205:110796.

Kochunas B, Huan X. Digital twin concepts with uncertainty for nuclear power applications. Energies. 2021;14(14):4235.

Huan X, Jagalur J, Marzouk Y. Optimal experimental design: Formulations and computations. Acta Numer. 2024;33:715-840.

Thelen A, Huan X, Paulson N, Onori S, Hu Z, Hu C. Probabilistic machine learning for battery health diagnostics and prognostics-review and perspectives. npj Mater Sustain. 2024;2(1):14.

Shin Y, Zhang Z, Karniadakis GE. Error estimates of residual minimization using neural networks for linear pdes. J Mach Learn Model Comput. 2023;4(4):73-98.

Hao Z, Zhang Z. Optimal regularity and error estimates of a spectral galerkin method for fractional advection-diffusion-reaction equations. SIAM J Numer Anal. 2020;58(1):211-33.

Jafarzadeh S, Silling S, Liu N, Zhang Z, Yu Y. Peridynamic neural operators: A data-driven nonlocal constitutive model for complex material responses. Comput Methods Appl Mech Eng. 2024;425:116914.

Yi J, Cheng J, Bessa MA. Practical multi-fidelity machine learning: Fusion of deterministic and bayesian models. arXiv preprint arXiv:2407.15110. 2024. Available from: https://arxiv.org/abs/2407.15110

Wang Q, Wu D, Li G, Liu Z, Tong J, Chen X, et al. Machine learning aided uncertainty quantification for engineering structures involving material-geometric randomness and data imperfection. Comput Methods Appl Mech Eng. 2024;423:116868.

Tian W, Li Q, Wang Q, Chen D, Gao W. Additive manufacturing error quantification on stability of composite sandwich plates with lattice-cores through machine learning technique. Compos Struct. 2024;327:117645.

Zhang M, Wang Q, Luo Z, Gao W. Virtual model-aided reliability analysis considering material and geometrical uncertainties for elastic metamaterials. Mech Syst Signal Process. 2024;211:111199.

D'Elia M, Du Q, Glusa C, Gunzburger M, Tian X, Zhou Z. Numerical methods for nonlocal and fractional models. Acta Numer. 2020;29:1-24.

Xu L, Abbaszadeh P, Moradkhani H, Chen N, Zhang X. Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index. Remote Sens Environ. 2020;250:112028.

Kovachki N, Lanthaler S, Mishra S. On universal approximation and error bounds for fourier neural operators. J Mach Learn Res. 2021;22(290):1-76.

Zhang J, Zheng Q, Wu L, Zeng L. Using deep learning to improve ensemble smoother: Applications to subsurface characterization. Water Resour Res. 2020;56(12):e2020WR027399.

Zhang J, Cao C, Nan T, Ju L, Zhou H, Zeng L. A novel deep learning approach for data assimilation of complex hydrological systems. Water Resour Res. 2024;60(2):e2023WR035389.

Cheng J, Ye Q, Lu Z, Zhang J, Zeng L, Parikh SJ, et al. Quantification of the sorption of organic pollutants to minerals via an improved mathematical model accounting for associations between minerals and soil organic matter. Environ Pollut. 2021;280:116991.

Xu W, Lyu C, Zhang J, Wang H, Wang R, Yan L, et al. Calibrating high-dimensional rock creep constitutive models for geological disaster prevention: An application of data assimilation methods. Int J Rock Mech Min Sci. 2024;183:105911.

Tran A, Wildey T, McCann S. Multi-fidelity machine-learning with uncertainty quantification and bayesian optimization for materials design: Application to ternary random alloys. Comput Mater Sci. 2020;181:109679.

Batra R, Song L, Ramprasad R. Emerging materials intelligence ecosystems propelled by machine learning. Nat Rev Mater. 2021;6(8):655-78.

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

Chen C, Zuo Y, Ye W, Li X, Ong SP. A critical review of machine learning of energy materials. Adv Energy Mater. 2020;10(8):1903242.

von Lilienfeld OA. Machine learning to design chemical systems. Angew Chem Int Ed. 2020;59(10):4166-71.

Olivetti EA, Cole JM, Kim E, Kononova O, Ceder G, Han TYJ, et al. Data-driven materials research enabled by natural language processing and information extraction. Appl Phys Rev. 2020;7(4):041317.

Fung V, Hu J, Ganesh P, Sumpter BG. Machine learned features from density of states for accurate adsorption energy prediction. Nat Commun. 2021;12(1):88.

Tran A, Furlan JM, Pagalthivarthi KV, Viswanathan MB, Wildey T. Explainable ai for material property prediction based on energy computation. Comput Mater Sci. 2023;215:111788.

Zhao H, Chen W. Uncertainty quantification and propagation in materials modeling using machine learning. Comput Mater Sci. 2022;207:111262.

Pernot P, Cailliez F. A critical review of statistical calibration methods for physical models. Mech Syst Signal Process. 2020;139:106620.

Zhang L, Lin DY, Wang H, Car R, E W. Active learning of potential energy surfaces with theoretical accuracy for materials modeling. Nat Mach Intell. 2020;2(2):66-73.

Pagani S, Fresca S, Manzone F, Deponti A, Quarteroni A. Enabling forward uncertainty quantification and sensitivity analysis in cardiac electrophysiology by reduced order modeling and machine learning. Int J Numer Method Biomed Eng. 2021;37(11):e3531.

Dan Y, Zhao Y, Li X, Li S, Hu M, Hu J. Generative adversarial networks based on wasserstein distance for knowledge graph embeddings. Knowl Based Syst. 2020;190:105165.

Yi J, Zhou Q, Cheng Y, Liu J. Efficient adaptive kriging-based reliability analysis combining new learning function and error-based stopping criterion. Struct Saf. 2020;83:101904.

Jin T, Chazirakis A, Kalligiannaki E, Harmandaris V, Katsoulakis MA. Data-driven uncertainty quantification for systematic coarse-grained models. Soft Mater. 2021;19(4):434-51.

Suzuki JL, Gulian M, Zayernouri M, D'Elia M. Fractional modeling in action: A survey of nonlocal models for subsurface transport, turbulent flows, and anomalous materials. J Peridyn Nonlocal Model. 2023;5(3):392-459.

Lykkegaard MB, Dodwell TJ, Moxey D. Accelerating uncertainty quantification of groundwater flow modelling using a deep neural network proxy. Comput Methods Appl Mech Eng. 2021;383:113895.