

ORIGINAL RESEARCH

Open access

# Algorithmic Consensus vs Scientific Consensus in Materials Prediction

Sanjay Kulkarni<sup>1</sup>, Meenal Joshi<sup>1\*</sup>, Rohan Patil<sup>2</sup>

## Abstract

In the rapidly evolving field of computational and data-driven materials engineering, the interplay between algorithmic processes and established scientific paradigms shapes the reliability of predictive outcomes. Traditional scientific consensus emerges from iterative experimental validation, peer review, and cumulative evidence, fostering a shared understanding of material behaviors and properties. In contrast, algorithmic consensus arises from the aggregation of computational models, often leveraging machine learning architectures to distill patterns from vast datasets. This manuscript explores the tensions and synergies between these two forms of consensus in materials prediction, highlighting how data-driven approaches can either reinforce or challenge longstanding scientific interpretations. A conceptual gap persists in integrating these consensus mechanisms, where algorithmic outputs may diverge from empirical benchmarks due to representation biases or uncertainty propagation. To address this, we introduce the Consensus Integration Lattice (CIL), a novel framework that structures the alignment of algorithmic and scientific consensus through layered computational workflows, feedback mechanisms, and epistemic risk assessments. By conceptualizing discovery pipelines that couple high-throughput simulations with multimodal data integration, CIL facilitates more robust materials predictions. Implications extend to autonomous discovery systems, inverse design strategies, and uncertainty quantification, potentially enhancing the efficiency of materials informatics ecosystems. This work underscores the need for infrastructure-level analyses to bridge computational agility with scientific rigor, paving the way for hybrid paradigms in materials engineering.

**Keywords** Materials informatics, Uncertainty quantification, Machine learning, Scientific consensus, Discovery pipelines, Algorithmic consensus

\*Correspondence:

Meenal Joshi  
meenal.joshi@gmail.com

<sup>1</sup> Department of Materials Data Science, Faculty of Engineering, Savitribai Phule Pune University, Pune, India

<sup>2</sup> Department of Computational Materials Systems, Faculty of Engineering, IIT Bombay, Mumbai, India

## Introduction

### The evolution of materials prediction paradigms

The field of materials science has undergone a profound epistemic and infrastructural transformation with the integration of computational and data-driven methodologies. Traditionally, materials discovery was anchored in experimentally intensive workflows guided by thermodynamic theory, quantum mechanics, and empirical heuristics. Foundational predictive insights emerged through iterative cycles of hypothesis formulation,

synthesis, characterization, and validation—a process both methodologically rigorous and temporally constrained [1, 2]. While this paradigm yielded robust scientific consensus, its throughput limitations restricted exploration across the vast combinatorial expanse of chemical and structural design spaces.

The emergence of high-performance computing and density functional theory (DFT) simulations marked the first major acceleration phase in predictive materials science. Computational thermodynamics enabled virtual screening of candidate compounds prior to experimental realization,

reducing cost and time barriers. However, it was the convergence of machine learning, large-scale databases, and automated workflows that catalyzed a second, more transformative paradigm shift [3, 4]. Data-driven materials engineering now operates within integrated ecosystems where predictive inference can precede mechanistic interpretation, reversing traditional epistemic sequences.

This transformation is institutionally exemplified by initiatives such as the Materials Genome Initiative, which formalized the role of computational infrastructure, open databases, and informatics pipelines in expediting materials innovation [5]. By coordinating simulation repositories, experimental datasets, and digital platforms, such infrastructures established the foundation for algorithmically mediated discovery.

Within this new paradigm, representation learning has emerged as a central predictive engine. Graph neural networks, crystal graph convolutions, and attention-based architectures encode atomic environments, symmetry operations, and bonding interactions into latent embeddings capable of forecasting functional properties [6, 7]. These systems now predict bandgaps, elastic moduli, catalytic activity, thermal transport coefficients, and phase stability with unprecedented speed. Importantly, their predictive logic is not deductive but statistical—arising from pattern extraction across structured datasets rather than direct derivation from first principles.

Yet, as predictive capacity expands, so too does epistemic complexity. Materials prediction is no longer a unidirectional translation of physical law into measurable outcome. Instead, it has evolved into a layered interpretive process where computational abstractions, training distributions, and optimization objectives shape the contours of predictive knowledge. Within this environment, consensus formation—historically rooted in experimental reproducibility—undergoes structural reconfiguration.

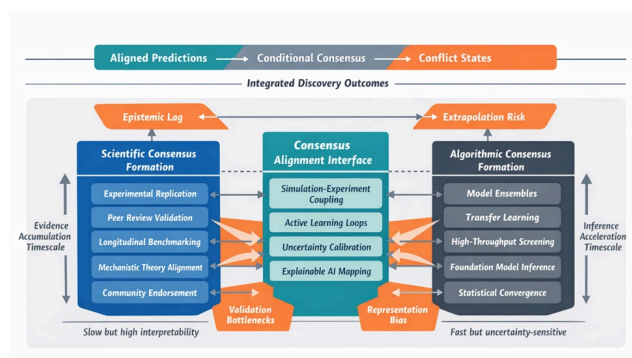
## Scientific consensus vs algorithmic consensus

Scientific consensus in materials science traditionally emerges through cumulative empirical validation. Experimental replication, cross-laboratory benchmarking, and longitudinal verification stabilize knowledge claims over time. Authoritative databases and review syntheses consolidate such evidence, producing community-endorsed reference points that guide subsequent discovery [8, 9].

Platforms such as the Materials Project and JARVIS function as infrastructural crystallizations of this consensus, aggregating validated datasets into standardized predictive baselines [10].

Algorithmic consensus, by contrast, is computationally emergent. It arises when ensembles of machine learning models converge upon similar predictions despite heterogeneity in architecture, training pathways, or representational encoding [11, 12]. Techniques such as transfer learning, multi-fidelity modeling, and federated training amplify this convergence, producing statistically reinforced outputs that simulate agreement.

However, this agreement is structurally distinct from scientific consensus. Whereas empirical consensus is grounded in physical validation, algorithmic consensus is rooted in distributional alignment. Models may agree not because predictions are physically correct, but because they are conditioned on similar datasets, biases, or optimization constraints. Consequently, high predictive accuracy within benchmark domains does not guarantee extrapolative reliability [13, 14]. These structurally distinct consensus pathways generate epistemic tensions in predictive interpretation (Figure 1).



**Figure 1.** Epistemic Divergence and Convergence Dynamics Between Algorithmic and Scientific Consensus

Comparative epistemic map illustrating divergence and convergence dynamics between algorithmic and scientific consensus in materials prediction. The figure contrasts computational agreement driven by model ensembles with empirically grounded validation processes, highlighting zones of alignment, epistemic lag, and risk amplification across discovery timelines.

This dual consensus structure introduces interpretive tensions. Algorithmic outputs can outpace experimental

verification cycles, generating rapid predictive frontiers that lack commensurate empirical grounding. In such contexts, consensus becomes temporally asynchronous: computational agreement precedes scientific validation. The resulting epistemic lag complicates decision-making in deployment-critical domains.

## Challenges in data-driven materials ecosystems

Central challenges within computational materials ecosystems stem from data heterogeneity and integration complexity. Contemporary discovery infrastructures synthesize multimodal inputs, including first-principles simulations, experimental measurements, process metadata, and literature-extracted knowledge graphs [15, 16]. While such multimodality enriches predictive scope, it also introduces representational incompatibilities and latent biases.

Data provenance, resolution disparities, and methodological noise can distort training distributions, embedding systematic skew into predictive architectures [17, 18]. Without robust harmonization protocols, models may overfit simulation artifacts or underrepresent experimental uncertainties. These distortions propagate through inference layers, shaping downstream consensus structures.

Uncertainty quantification thus emerges as a critical stabilizing mechanism. Epistemic uncertainty—arising from model ignorance—and aleatoric uncertainty—stemming from data variability—must be disentangled to prevent overconfident predictions [19, 20]. Failure to operationalize uncertainty within decision pipelines risks reinforcing algorithmic consensus even when predictive reliability is low.

Inverse design workflows further intensify these challenges. Here, models do not merely predict properties but actively steer compositional exploration toward target performance metrics [17, 21]. Optimization pressures may privilege computational tractability over physical plausibility, generating candidate materials that satisfy algorithmic criteria while remaining experimentally infeasible. This divergence underscores a growing epistemic disconnect between predictive inference and scientific validation.

Active learning and closed-loop experimentation partially address these issues by iteratively integrating experimental

feedback into model retraining cycles [22, 23]. Yet, such systems often optimize efficiency rather than consensus coherence. Formal mechanisms for aligning algorithmic agreement with scientific verification remain underdeveloped.

## Scalability, foundation models, and interpretability trade-offs

The scaling of machine learning infrastructures introduces additional consensus complexities. Foundation models trained on expansive materials corpora promise cross-domain generalizability, enabling transfer across property classes and compositional families [24, 25]. Their latent embeddings capture high-order relational structures inaccessible to smaller architectures.

However, this scaling amplifies interpretability challenges. Black-box inference obscures mechanistic reasoning pathways, complicating efforts to reconcile predictions with established physical theory. While explainable AI techniques offer partial transparency, they often operate post hoc, interpreting rather than governing predictive logic.

High-throughput screening platforms further exacerbate this trade-off. Millions of candidate materials can be computationally evaluated within compressed timeframes [4, 6]. Yet predictive volume does not equate to epistemic depth. Without structured validation hierarchies, rapid screening risks generating expansive yet shallow knowledge landscapes—broad in scope but fragile in reliability.

Thus, computational scalability introduces an infrastructural paradox: the very systems that accelerate discovery also strain the mechanisms required to legitimize it.

## Bridging computational and epistemic dimensions

Addressing these tensions necessitates a systems-level reconceptualization of materials prediction. Rather than viewing prediction as an isolated modeling task, it must be understood as a layered ecosystem encompassing representation, inference, validation, and consensus formation [26, 27]. Each layer contributes distinct epistemic signals that collectively shape discovery trajectories.

Emerging advances in autonomous experimentation, robotic synthesis, and reinforcement learning offer dynamic

coupling between computational inference and empirical validation [28, 29]. These infrastructures enable adaptive discovery loops where predictive hypotheses are experimentally tested and recursively refined. However, while such systems enhance feedback velocity, they do not inherently resolve consensus misalignment.

The distinction between algorithmic and scientific consensus becomes particularly consequential in high-stakes application domains. Energy storage materials, carbon capture catalysts, quantum semiconductors, and biomedical alloys demand predictive reliability that extends beyond statistical performance [9, 30]. Deployment risks—economic, environmental, and societal—necessitate consensus structures that are not only computationally robust but epistemically coherent.

## Positioning the Consensus Integration Lattice (CIL)

In response to these structural challenges, this manuscript introduces the Consensus Integration Lattice (CIL) as a novel interpretive and infrastructural framework for materials prediction ecosystems. CIL conceptualizes consensus not as an emergent by-product of modeling convergence but as an architected outcome shaped by data lineage, representational fidelity, uncertainty propagation, and validation coupling.

By structuring interactions across data, model, and discovery layers through lattice-like feedback architectures, the framework provides computational steering logics that reconcile rapid algorithmic inference with deliberative scientific verification. Rather than privileging one consensus mode over the other, CIL operationalizes their integration—aligning predictive scalability with epistemic accountability.

Through this lens, materials prediction is reframed as a consensus-engineered process, where discovery acceleration and knowledge reliability co-evolve within structured computational ecosystems.

## Theoretical Background & Literature Synthesis

### Foundations of computational materials informatics

Materials informatics has emerged as a cornerstone of modern materials engineering, integrating data science principles with domain-specific knowledge to accelerate discovery [1, 2]. At its core, this field relies on machine learning to extract actionable insights from complex datasets, encompassing atomic structures, electronic properties, and synthesis parameters [3, 4]. Representation learning, particularly through graph neural networks, transforms raw material descriptors into latent spaces conducive to prediction tasks [6, 7]. These representations enable models to capture topological and relational features, facilitating applications in property forecasting and structure-property mapping [11, 12].

High-throughput computation complements these efforts by generating extensive simulation data, often via density functional theory or molecular dynamics, which populate databases for training [5, 10]. Such infrastructures support the Materials Genome Initiative's vision of reducing development timelines, but they also introduce dependencies on data quality and model robustness [8, 9]. Literature emphasizes the role of small data regimes, where transfer learning mitigates scarcity by leveraging pre-trained models from related domains [3, 7].

### Machine learning architectures and uncertainty in prediction

Deep learning architectures, including convolutional and recurrent variants adapted for materials, have revolutionized prediction accuracy [2, 4]. Graph-based models, in particular, excel in handling crystalline and molecular graphs, enabling end-to-end learning from composition to performance [6, 14]. However, challenges in generalizability persist, as models trained on specific datasets may underperform on out-of-distribution samples [13, 16]. Uncertainty quantification addresses this by estimating prediction confidence, often through Bayesian frameworks or ensemble techniques, which inform decision-making in discovery workflows [15, 19].

Explainable AI techniques further dissect model decisions, revealing how features contribute to outputs and aligning them with physical intuitions [11, 17]. In materials contexts, this transparency aids in identifying biases, such as those arising from imbalanced datasets or incomplete representations [13, 18]. Literature synthesizes these elements into pipelines that couple simulation with experimentation, ensuring iterative refinement [20, 23].

## Autonomous and closed-loop discovery systems

Autonomous discovery systems represent a maturation of data-driven approaches, incorporating robotics and AI for self-guided experimentation [22, 23]. Closed-loop paradigms integrate prediction, synthesis, and characterization in real-time, optimizing parameters via active learning [15, 19]. Feedback mechanisms, such as Bayesian optimization, steer explorations toward high-value regions in material space [21, 28]. These systems exemplify algorithmic consensus in action, where iterative model updates converge on optimal designs [17, 29].

Inverse materials design flips traditional workflows, starting from desired properties to infer compositions, often using generative models or reinforcement learning [17, 21]. Multimodal datasets enhance this by fusing disparate sources, like spectroscopic data with computational simulations, to enrich feature spaces [24, 25]. Yet, epistemic risks emerge when algorithmic convergence outpaces scientific validation, potentially leading to overconfidence in unverified predictions [13, 30].

## Integration of simulation-experiment coupling and foundation models

Simulation-experiment coupling bridges computational predictions with empirical reality, using machine learning to calibrate models against experimental outcomes [16, 20]. This hybrid approach mitigates discrepancies, fostering a more unified consensus [5, 9]. Foundation models for science, pre-trained on broad scientific corpora, offer versatile starting points for materials tasks, adapting via fine-tuning to specific predictions [24, 25].

Literature highlights trade-offs in these integrations: while coupling enhances reliability, it demands sophisticated uncertainty handling to avoid error amplification [19, 26]. In high-stakes domains like energy materials, such dynamics underscore the need for epistemic structures that evaluate consensus quality [9, 31].

## Epistemic and infrastructure trade-offs in consensus formation

The synthesis of these elements reveals underlying trade-offs in materials prediction ecosystems. Algorithmic consensus, driven by model ensembles and data

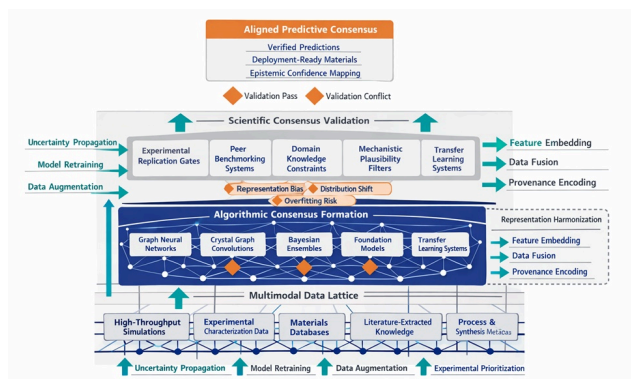
aggregation, prioritizes efficiency and scalability [10, 12]. Scientific consensus, conversely, emphasizes reproducibility and peer scrutiny, often slower but more resilient to anomalies [8, 27]. Disparities arise in representation-inference interactions, where data biases can skew algorithmic outputs away from scientific norms [6, 14].

Infrastructure-level analyses advocate for discovery steering logics that incorporate both, such as adaptive sampling to balance exploration and exploitation [15, 22]. Feedback loops in autonomous systems exemplify this, dynamically adjusting based on uncertainty metrics [23, 28]. Overall, the literature points to a need for frameworks that interpret these interactions, ensuring computational workflows align with epistemic goals [29, 30]. Key infrastructural and epistemic distinctions between consensus modes are synthesized in **Table 1**.

**Table 1.** Comparative Infrastructure and Epistemic Characteristics of Algorithmic vs Scientific Consensus in Materials Prediction

Dimension	Algorithmic Consensus	Scientific Consensus	Integrated Implications (CIL Level)
Formation Mechanism	Model ensembles, statistical convergence	Experimental replication, peer validation	Requires lattice-mediated alignment
Timescale	Rapid, computation-driven	Slow, evidence-driven	Temporal synchronization needed
Data Dependency	Training distributions	Empirical observations	Provenance harmonization critical
Interpretability	Often opaque / black-box	Mechanistically grounded	Explainability bridges gaps
Uncertainty Handling	Probabilistic estimation	Experimental variance analysis	Unified uncertainty propagation
Scalability	Extremely high	Resource-limited	Hybrid screening hierarchies

Failure Modes	Overfitting, bias amplification	Measurement error, reproducibility limits	Cross validation mitigate
Validation Logic	Benchmark accuracy	Physical plausibility	Dual-gate validation
Infrastructure Anchor	HPC + ML platforms	Laboratories + characterization systems	Coupled computational physics ecosystem
Consensus Robustness	Distribution-sensitive	Evidence-resilient	Weighted integration requirements



**Figure 2.** Consensus Integration Lattice (CIL): Layered Architecture for Aligning Algorithmic and Scientific Consensus in Materials Prediction

## Proposed conceptual framework

### Overview of the Consensus Integration Lattice (CIL)

The Consensus Integration Lattice (CIL) is introduced as an original framework to conceptualize the alignment between algorithmic and scientific consensus in materials prediction. CIL structures this alignment through a multi-layered architecture that integrates data ingestion, model orchestration, and discovery validation, emphasizing computational steering logics to navigate epistemic risks. At its base layer, CIL posits a data lattice where multimodal inputs—spanning simulations, experiments, and informatics repositories—are woven into a unified representation space. This layer facilitates the transition from raw data to feature embeddings, enabling models to capture relational dynamics without empirical overfitting.

Ascending the lattice, the model orchestration layer aggregates diverse architectures, such as graph neural networks and Bayesian ensembles, to form preliminary algorithmic consensus. Feedback loops recirculate discrepancies back to data refinement, ensuring iterative coherence. The apex layer incorporates scientific validation gates, where predictions are filtered through epistemic risk structures derived from uncertainty mappings and domain knowledge integration. CIL's pipelines thus flow from data to model to discovery, with bidirectional arrows representing dynamic adjustments (Figure 2).

Conceptual architecture of the Consensus Integration Lattice (CIL) illustrating the multi-layered integration of algorithmic and scientific consensus across data, model, and validation infrastructures. The lattice structure visualizes multimodal data ingestion, ensemble model orchestration, and epistemic validation gates, interconnected through uncertainty-mediated feedback loops. Bidirectional steering pathways depict how consensus divergences propagate and are resolved across discovery workflows.

### Key dynamics and computational steering

Central to CIL are the steering logics that guide consensus formation. These logics operate through representation-inference interactions, where data embeddings inform model predictions, and inference outputs refine representations. Trade-offs are managed by prioritizing epistemic coherence over raw predictive speed, avoiding the pitfalls of isolated algorithmic convergence.

One dynamic can be conceptualized as the consensus

alignment function, expressed as  $C = \frac{\alpha A}{\alpha A + (1 - \alpha) S}$ , where C represents integrated consensus, A algorithmic contributions from model ensembles, S scientific benchmarks, and  $\alpha$  a weighting factor modulated by uncertainty levels. This captures the interaction between computational agility and empirical grounding, allowing for adaptive blending based on discovery context.

Another aspect involves feedback efficiency, which may be expressed as  $F = \frac{\Delta U}{\Delta I}$ , with F denoting feedback strength,  $\Delta U$  change in uncertainty, and  $\Delta I$  iterative inputs.

This formalizes how loops reduce epistemic risks by quantifying adjustments in prediction workflows.

Finally, the discovery steering trade-off is captured as  $T = \log(EV)$  where T is the trade-off metric, E exploration breadth, and V validation depth. This logarithmic form emphasizes the non-linear balance in expanding search spaces while maintaining scientific fidelity.

These formulas underscore CIL's interpretive power, framing materials prediction as a lattice of interdependent processes rather than linear pipelines.

## Algorithmic consensus vs scientific consensus in materials prediction

### Analytical implications

The Consensus Integration Lattice (CIL) offers a lens through which to examine the broader implications for computational workflows in materials prediction, emphasizing how layered structures can mitigate divergences between algorithmic and scientific consensus. In data ingestion phases, CIL's base layer implies a need for enhanced representation strategies that accommodate multimodal inputs, reducing biases that often lead to algorithmic overconfidence [6, 7]. This has direct bearing on materials informatics, where incomplete datasets can skew predictions; by integrating feedback from higher layers, CIL suggests pathways for dynamic data curation, ensuring that representations evolve in tandem with inference demands [1, 3].

In model orchestration, the framework highlights trade-offs in architecture selection, where ensemble methods foster algorithmic consensus but may dilute scientific interpretability if not anchored by uncertainty metrics [11, 12]. For instance, in high-throughput screening, CIL's steering logics imply that model diversity should be balanced against convergence speed, preventing premature consensus that overlooks epistemic risks [4, 10]. This analytical perspective extends to inverse design, where target-driven searches benefit from lattice-mediated adjustments, aligning computational explorations with physical constraints [17, 21].

Discovery pipelines under CIL reveal implications for autonomous systems, where closed-loop mechanisms can be steered to incorporate scientific validation at key junctures [22, 23]. The framework's feedback loops imply a

reduction in iterative overhead by prioritizing high-uncertainty regions, thus optimizing resource allocation in coupled simulation-experiment setups [16, 20]. Epistemic risk structures within CIL further imply a systematic approach to quantifying consensus quality, potentially informing infrastructure designs that embed risk assessments into real-time workflows [13, 19].

One implication can be formalized as the risk propagation

dynamic, expressed as  $R = \sum_i^n u_i w_i$ , where R is aggregate

epistemic risk,  $u_i$  uncertainty from layer i, and  $w_i$  weights reflecting interlayer dependencies. This captures the cumulative interaction of uncertainties across the lattice, guiding computational adjustments to minimize overall risk.

Another formalization addresses workflow efficiency, which

may be expressed as  $E = \frac{PC}{F}$ , with E efficiency, P

predictive fidelity, C computational cost, and F feedback iterations. This equation interprets the balance in discovery steering, highlighting how CIL's logics can enhance throughput without sacrificing consensus integrity.

A third dynamic involves consensus robustness,

conceptualized as  $B = \exp(-\beta(A - S)^2)$ , where B is

robustness,  $\beta$  a sensitivity parameter, and  $\frac{A}{S}$  the divergence between algorithmic and scientific components. This Gaussian-like form underscores the exponential decay in reliability as divergences grow, implying tighter integration via lattice mechanisms.

These implications collectively steer towards hybrid infrastructures, where CIL facilitates the coupling of foundation models with domain-specific validations, enhancing generalizability in materials ecosystems [24, 25]. In uncertainty quantification, the framework implies advanced propagation models that account for lattice interactions, potentially reducing errors in property predictions [14, 15]. For representation learning, CIL's layers suggest adaptive embeddings that respond to consensus feedback, improving transferability across material classes [2, 26].

Overall, these analytical insights position CIL as a tool for infrastructure-level optimization, where trade-offs in speed, accuracy, and verifiability are navigated through interpretive

structures [5, 9]. This extends to sustainable materials design, implying accelerated pipelines that maintain scientific rigor [30, 31].

## Results and Discussion

The introduction of the Consensus Integration Lattice (CIL) prompts a reevaluation of how consensus mechanisms operate within computational and data-driven materials engineering, revealing both synergies and persistent challenges. CIL's layered approach underscores the potential for algorithmic consensus to augment scientific processes, particularly in scenarios where data volume outpaces traditional validation [1, 4]. However, this integration is not without friction; representation biases in base layers can amplify divergences, necessitating refined data-model interactions to align outputs with empirical realities [6, 13].

In the context of machine learning architectures, CIL highlights the value of ensemble strategies in building robust algorithmic consensus, yet it also exposes limitations in black-box models that hinder scientific interpretability [11, 12]. Discussion here centers on the need for explainable frameworks that bridge these gaps, allowing for computational steering that incorporates domain knowledge without compromising agility [17, 24]. High-throughput and autonomous systems exemplify this, where CIL's feedback loops could enhance closed-loop efficiency, but require careful calibration to avoid over-reliance on algorithmic signals [22, 23, 28].

Epistemic risk structures within CIL invite discourse on uncertainty management, where propagation across layers demands sophisticated quantification to prevent consensus erosion [15, 19]. This is particularly relevant in inverse design and multimodal integration, where CIL implies a shift towards adaptive workflows that dynamically adjust to risk profiles [18, 20, 21]. Yet, challenges arise in scaling these to real-world infrastructures, as computational costs may offset gains in discovery speed [5, 10].

Broader field dynamics suggest that CIL could inform policy in materials informatics, promoting standards for consensus evaluation that transcend individual models [8, 9]. For instance, in foundation models and simulation coupling, the framework encourages hybrid paradigms that leverage pre-trained generality while enforcing scientific anchors [16, 25]. Trade-offs in exploration versus validation, as

formalized in CIL, further discuss the balance required for innovative yet reliable predictions [26, 29, 31].

Limitations of CIL itself warrant attention: its conceptual nature assumes ideal interoperability, which may not hold in fragmented ecosystems [3, 14]. Future extensions could incorporate real-time data streams, enhancing its applicability to fast-evolving domains like energy materials [27, 30]. Ultimately, CIL fosters a discourse on epistemic coherence, urging the community towards integrated systems that harmonize computational and scientific strengths [2, 7].

## Conclusion

The exploration of algorithmic versus scientific consensus in materials prediction illuminates critical junctures where computational agility meets empirical rigor, with the Consensus Integration Lattice (CIL) serving as a unifying framework. By structuring data-model-discovery pipelines through layers, feedback, and steering logics, CIL provides interpretive tools to navigate consensus divergences, enhancing the reliability of predictive outcomes in materials engineering.

Key insights from this manuscript emphasize the implications for workflows, where epistemic risk assessments and representation-inference dynamics foster more cohesive integrations. CIL's formalizations of dynamics like risk propagation and efficiency trade-offs offer conceptual handles for optimizing infrastructures, potentially accelerating discoveries in autonomous and inverse design contexts.

Looking forward, CIL invites further development in hybrid systems, where algorithmic consensus complements scientific validation to address complex challenges in materials informatics. This work advocates for continued emphasis on computational steering to bridge paradigms, ultimately advancing data-driven materials science towards more robust and interpretable predictions.

## Acknowledgements

None

## Conflict of interest

None

## Financial support

None

None

Received: 30 Mar 2023 Revised: 26 Jul 2023 Accepted: 03 Oct 2023  
Published online: 18 March 2024

## Ethics statement

### Rights and permissions

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

## References

- Ramprasad R, Batra R, Pilania G, Mannodi-Kanakithodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *npj Comput Mater.* 2017;3(1):54.
- Fung V, Hu G, Ganesh P, Sumpter BG. Recent advances and applications of deep learning methods in materials science. *npj Comput Mater.* 2022;8(1):72.
- Yu K, Li H, Li Z. Small data machine learning in materials science. *npj Comput Mater.* 2023;9(1):100.
- Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. *npj Comput Mater.* 2019;5(1):83.
- de Pablo JJ, Jackson NE, Webb MA, Chen L-Q, Moore JE, Morgan D, et al. New frontiers for the materials genome initiative. *npj Comput Mater.* 2019;5(1):173.
- Duan K, Tan Y, Chen J, Gu X, Li W, Chen Z, et al. MaterialsAtlas.org: A materials informatics web app platform for materials discovery and survey of state-of-the-art. *npj Comput Mater.* 2022;8(1):75.
- Buterez D, Janet JP, Kiddle SJ, Oglic D, Liò P. Transfer learning with graph neural networks for improved molecular property prediction in the multi-fidelity setting. *Nat Commun.* 2024;15(1):5566.
- Li J, Xue D, Li Z. The mastery of details in the workflow of materials machine learning. *npj Comput Mater.* 2024;10(1):331.
- Wang J, Chen H, Hu Z, Yoo M, Hu Q, Guo H, et al. Machine learning for a sustainable energy future. *Nat Rev Mater.* 2023;8(2):90-105.
- Choudhary K, Wines D. JARVIS-Leaderboard: a large scale benchmark of materials design methods. *npj Comput Mater.* 2024;10(1):97.
- Dai Y, Hu G, Ganesh P, Sumpter BG. Explainable machine learning in materials science. *npj Comput Mater.* 2022;8(1):84.
- Singh K, Batra R, Pilania G, Ramprasad R. Exploiting redundancy in large materials datasets for efficient machine learning with less data. *Nat Commun.* 2023;14(1):7292.
- Chen C, Zuo Y, Ye W, Li X, Ong SP. A critical examination of robustness and generalizability of machine learning prediction of materials properties. *npj Comput Mater.* 2023;9(1):101.
- Wei Y, Yang R, Hou Y, Ramprasad R, Gao H. Machine learning in concrete science: applications, challenges, and best practices. *npj Comput Mater.* 2022;8(1):810.
- Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater.* 2019;5(1):21.
- Maksov A, Ziatdinov M, Fujii K, Sumpter B, Kalinin SV. Ensemble learning-iterative training machine learning for uncertainty quantification and automated experiment in atom-resolved microscopy. *npj Comput Mater.* 2021;7(1):69.
- Karpovich C, Pan E, Olivetti EA. Deep reinforcement learning for inverse inorganic materials design. *npj Comput Mater.* 2024;10(1):287.
- Li J, Chen Y, Qian C, Liu R, Ma T, Jing M, et al. Artificial-intelligence-led revolution of construction materials: From

molecules to Industry 4.0. *Matter*. 2023;6(6):1831-59.

Kusne AG, McDannald G. Targeted materials discovery using Bayesian algorithm execution. *npj Comput Mater*. 2024;10(1):326.

Kusne AG, McDannald G, DeCost B, Oses C, Toher C, Curtarolo S, et al. Closed-Loop error-correction learning accelerates experimental discovery of thermoelectric materials. *Adv Mater*. 2023;35(30):2302575.

Kanakithodi AKM, Pilania G, Ramprasad R. Attribute driven inverse materials design using deep learning Bayesian framework. *npj Comput Mater*. 2019;5(1):263.

Priyadarshini K, Abolhasani M, Bateni F, Coley CW, Epps RW, Li Y, et al. Performance metrics to unleash the power of self-driving labs in chemistry and materials science. *Nat Commun*. 2024;15(1):5569.

Epps RW, Volk AA, Reyes KG, Abolhasani M. Two-step machine learning enables optimized nanoparticle synthesis. *npj Comput Mater*. 2021;7(1):20.

Chen C, Gu GX. Artificial Intelligence to Power the Future of Materials Science and Engineering. *Adv Intell Syst*. 2020;2(1):1900143.

Kuenneth C, Ramprasad R. AI Applications through the Whole Life Cycle of Material Discovery. *Matter*. 2020;3(2):393-432.

Lee K, Chen Y, Li K, Batra R, Pilania G, Ramprasad R. Data-Driven design for metamaterials and multiscale systems: A review. *Adv Mater*. 2024;36(3):2305254.

Simonetti M, Boyd PG, Smit B. Computational modeling of reticular materials: The past, the present, and the future. *Adv Mater*. 2024;36(41):2412005.

Stach E, DeCost B, Kusne AG, Hattrick-Simpers J, Brown KA, Reyes KG, et al. Autonomous experimentation systems for materials development: A community perspective. *Matter*. 2021;4(9):2702-26.

Clark K, Tian Y, Oyola-Reynoso S, Abel B, VahidMohammadi A, Razal JM, et al. Physical computing for materials acceleration platforms. *Matter*. 2022;5(10):3314-32.

Kusne AG, McDannald G. Flexible formulation of value for experiment interpretation and design. *Matter*. 2024;7(2):803-10.

Leo J-P, Sun Y, Geng S, Li L, Singh CV, Peng C. Machine intelligence accelerated design of conductive MXene aerogels with programmable properties. *Nat Commun*. 2024;15(1):9011.