ORIGINAL RESEARCH **Open access**

Simulation Priors in Machine Learning Materials Models: Hidden Physics Assumptions

Ahmed Mansour[1*], Omar Saeed[1]

Abstract

The integration of machine learning into materials engineering has transformed discovery pipelines by leveraging vast simulation-generated datasets and high-throughput computational workflows. Within this data-driven paradigm, models frequently incorporate simulation priors—implicit assumptions derived from physical approximations, boundary conditions, and discretization choices embedded in first-principles calculations or molecular dynamics trajectories. These priors, often hidden within representation learning and graph-based architectures, introduce epistemic biases that propagate through inference to downstream tasks such as inverse design and closed-loop experimentation. A key conceptual gap lies in the lack of systematic frameworks for articulating and managing these assumptions as integral components of the computational infrastructure rather than incidental data artifacts. This article introduces the Simulation Prior Articulation Framework (SPAF), an original systems-level conceptual structure that delineates layered processing of multimodal materials data, explicit prior extraction from simulation ecosystems, integration into deep learning architectures, and steering of discovery pipelines via feedback mechanisms. SPAF emphasizes representation–inference interactions, computational workflow dynamics, and infrastructure trade-offs to enhance simulation–experiment coupling without empirical benchmarking. By framing hidden physics assumptions as addressable epistemic structures, the framework provides integrative insights for materials informatics, foundation models, and autonomous discovery systems, supporting more transparent and robust data-driven materials engineering pipelines.

Keywords Data-driven discovery, Graph neural networks, Inverse materials design, Simulation priors, Machine learning materials models, Hidden physics assumptions

*Correspondence:
Ahmed Mansour
ahmed.mansour@gmail.com
[1] Department of Materials Engineering and Data Modeling, Faculty of Engineering, Cairo University, Cairo, Egypt

Introduction

The emergence of machine learning in materials discovery

Over the past decade, machine learning has transitioned from a peripheral analytical aid to a central epistemic engine within computational materials science. This transformation has been catalyzed by the convergence of high-throughput quantum-mechanical simulations, large-scale data infrastructures, and advances in deep learning architectures capable of extracting structure–property relationships from high-dimensional materials spaces. Machine learning models now routinely ingest outputs from density functional theory (DFT), molecular dynamics (MD), and phase-field simulations, enabling accelerated prediction of thermodynamic stability, electronic structure, mechanical response, and transport properties at scales previously inaccessible through conventional computation alone [1, 2].

A defining feature of this paradigm is the rise of representation learning—the automated construction of latent descriptors that encode materials structures in forms amenable to statistical inference. Graph neural networks (GNNs), crystal graph convolutional networks, and message-passing architectures have been particularly influential, enabling atomistic systems to be represented as relational graphs in which nodes correspond to atomic species and edges encode bonding environments or spatial proximities [3-5]. These representations transcend handcrafted descriptors by capturing emergent interaction patterns derived from simulation outputs, allowing models to generalize across compositional and structural variations.

The integration of such architectures into high-throughput screening workflows has reshaped discovery logics. Rather than sequential hypothesis testing, materials exploration increasingly operates through parallelized evaluation of vast candidate spaces, guided by predictive models that prioritize promising chemistries and structures. This computational acceleration has supported the emergence of autonomous discovery systems—closed-loop platforms that couple simulation, machine learning, and robotic experimentation into iterative optimization cycles [6-8]. Within these ecosystems, machine learning does not merely interpret data; it actively steers experimental and computational trajectories, influencing which materials are synthesized, simulated, or discarded.

The role of simulation in data generation for materials models

Despite the centrality of machine learning, simulation infrastructures remain the foundational data engines that sustain contemporary materials AI. Quantum-mechanical calculations, atomistic simulations, and mesoscale modeling pipelines generate the bulk of structured datasets used to train predictive architectures. Among these, DFT has become the dominant backbone for electronic, structural, and energetic property prediction, while molecular dynamics provides temporal evolution insights into diffusion, phase transitions, and defect behavior. High-throughput frameworks operationalize these methods at scale, producing repositories containing millions of computed structures and associated properties [9, 10].

These datasets are increasingly multimodal. Simulation outputs are now fused with experimental measurements, spectroscopic signatures, microstructural imaging, and thermodynamic assessments, forming hybrid data ecosystems that enrich model training and validation [11, 12]. Multimodal fusion enhances predictive robustness by embedding complementary perspectives—linking idealized computational predictions with empirically observed phenomena.

However, simulation pipelines are not neutral data generators. Each computational step encodes methodological decisions and approximations that shape the resulting data distributions. Exchange–correlation functional selection in DFT, pseudopotential parameterization, basis set truncation, harmonic approximations in phonon calculations, timestep discretization in MD, and ensemble selection all introduce embedded physics assumptions. These assumptions function as latent priors—structural constraints that influence calculated energies, structural relaxations, and derived properties. When aggregated across high-throughput infrastructures, such priors become statistically entrenched, shaping the epistemic boundaries within which machine learning models operate. Common sources of simulation priors and the mechanisms by which they propagate into learned representations and predictions are summarized in **Table 1.**

Table 1. Taxonomy of simulation priors in materials datasets and their typical downstream consequences in machine learning pipelines

Simulation prior class	Where it enters the pipeline	Typical “hidden” assumption	How it propagates into ML representations	Likely failure mode if unarticulated	Example mitigation handle within SPAF

Exchange–correlation (XC) choice (DFT)	Electronic-structure calculation	Functional-specific bias (e.g., over/under-binding)	Shifts energy/forces distributions → embedding drift	Systematic error under extrapolation; misleading stability ranking	Prior token + validity domain tags; post-hoc calibration hooks
Pseudopotentials / basis sets	DFT setup	Approximate core treatment / truncation	Alters local environments learned by message passing	Inconsistent transfer across element families	Prior registry with provenance + compatibility constraints
Boundary conditions / cell choices	DFT/MD cell definition	Periodicity and finite-size representativity	Encodes artificial symmetry or suppressed defects	Poor generalization to surfaces, interfaces, defects	Prior articulation + scope flags (bulk vs surface vs defect)
Discretization / convergence thresholds	DFT/MD numerical settings	“Good enough” numerical stability	Introduces structured noise correlated with material class	Spurious correlations; unstable uncertainty estimates	Metadata enrichment + uncertainty tagging (numerical vs physical)
Ensemble & thermostat/barostat (MD)	MD trajectories	Stationarity and equilibrium assumptions	Latent dynamics bias in trajectory embeddings	Wrong kinetics/diffusion under different conditions	Prior strength parameter + mismatch detector in steering layer
Harmonic approximations (phonons)	Vibrational property computation	Neglect of anharmonicity	Temperature-related embedding misalignment	Breakdown at high T; wrong stability of phases	Prior scope constraints + targeted high-fidelity enrichment loop
Dataset curation / filtering rules	Post-processing	Removal of “outliers” as errors	Narrows representation support	Epistemic narrowing; blind spots	Prior extraction from curation logs + steering against coverage collapse

Challenges arising from hidden physics assumptions

The implicit nature of these simulation priors introduces a series of systemic challenges for data-driven discovery. When physics assumptions remain unarticulated, they become difficult to interrogate, quantify, or correct. Their influence often surfaces only when models are deployed beyond the regimes in which training data were generated. Under extrapolative conditions—novel chemistries, extreme thermodynamic environments, or non-equilibrium structures—prediction reliability can degrade as latent simulation constraints misalign with real-world physics [13-15].

In representation learning contexts, these issues can propagate and amplify. Graph-based encodings derived from simulation outputs may embed artifacts associated with specific computational settings—such as overbinding tendencies of certain functionals or suppressed anharmonicity in harmonic approximations. As a result, learned embeddings risk conflating physical signal with simulation artifact, influencing downstream predictions and uncertainty estimates [16-18]. This becomes particularly consequential in inverse design workflows, where generative or optimization models propose

candidate materials based on learned structure–property relationships. If those relationships reflect simulation-specific distortions, proposed materials may exhibit degraded performance when synthesized or experimentally characterized.

Closed-loop discovery platforms intensify this dynamic. Autonomous systems iteratively retrain models on newly generated simulation data, reinforcing existing priors through feedback cycles. Foundation models trained on aggregated materials databases similarly inherit cumulative biases embedded across constituent datasets [19-21]. Over successive iterations, these infrastructures risk epistemic narrowing—converging toward discovery pathways constrained by unexamined simulation assumptions.

At an infrastructural scale, this challenge manifests as a trade-off between simulation fidelity and data volume. High-fidelity simulations incorporating advanced functionals, large supercells, or explicit temperature effects yield more physically representative data but at reduced throughput. Conversely, lower-fidelity approximations enable expansive datasets but embed stronger simplifying assumptions. When such trade-offs remain unarticulated, machine learning systems inherit epistemic risk—optimizing predictions within bounded simulation realities rather than broader materials possibility spaces [22-24].

Positioning of the current work

Scholarly discourse has begun to address adjacent dimensions of this problem space. Advances in explainable machine learning seek to render model decision pathways interpretable, illuminating feature importance and representational saliency [25-27]. Physically informed neural networks integrate governing equations and conservation laws to constrain learning processes, embedding physics directly into model architectures [28, 29]. Autonomous experimentation frameworks further attempt to reconcile computational predictions with empirical validation through iterative lab–model coupling [30].

While these developments enhance interpretability, physical consistency, and experimental integration, they seldom treat simulation priors themselves as first-class analytical objects. The physics assumptions embedded in data generation pipelines remain diffusely distributed across computational workflows, rarely formalized within discovery architectures. As a result, their epistemic influence persists largely ungoverned—shaping model behavior, uncertainty landscapes, and exploration trajectories without explicit articulation.

The Simulation Prior Articulation Framework (SPAF) is introduced to address this conceptual gap. SPAF reconceptualizes simulation assumptions as steerable infrastructural elements rather than static background conditions. The framework establishes a layered architecture spanning data generation, representation learning, model inference, and discovery steering, embedding feedback channels that surface, quantify, and adapt simulation priors across iterative cycles.

By formalizing hidden physics assumptions within a systems-level structure, SPAF enables new analytical vantage points on materials AI ecosystems. It supports interrogation of how priors propagate through representation spaces, how they interact with uncertainty quantification, and how they influence inverse design and autonomous experimentation. More broadly, SPAF reframes simulation not merely as a data provider but as an epistemic actor within discovery systems—one whose embedded assumptions can be articulated, steered, and optimized in alignment with broader scientific objectives.

Theoretical Background & Literature Synthesis

Materials informatics and representation learning

Materials informatics has matured into a foundational pillar of computational materials science, driven by the systematic integration of machine learning into data-rich discovery infrastructures. Early developments positioned machine learning primarily as a predictive layer operating on curated structural descriptors—composition vectors, symmetry functions, and

thermodynamic parameters derived from simulation datasets [2]. These early predictive models demonstrated that statistical inference could approximate structure–property relationships with significant computational savings relative to first-principles simulations [1]. However, their dependence on handcrafted descriptors limited generalizability, particularly when extrapolating to unexplored chemistries or structural motifs.

Representation learning transformed this landscape by shifting descriptor construction from manual engineering to automated latent encoding. Deep learning architectures began to generate embeddings directly from raw structural inputs, learning hierarchical representations that captured both local atomic environments and extended crystalline periodicities. Critically, these embeddings were designed to respect materials-specific invariances—translational symmetry, rotational equivariance, and permutation invariance—ensuring physical plausibility in learned representations [29, 30].

Graph-based encodings emerged as the dominant representational paradigm within this shift. By modeling atoms as nodes and interatomic interactions as edges, graph representations translate simulation-derived structures into relational data formats suitable for deep learning. Edge features may encode bond lengths, coordination environments, or interaction energies derived from simulation outputs, embedding physics-informed relationality within the representation space [3, 4]. These encodings enable models to operate across variable-sized systems without dimensionality constraints, facilitating scalable learning across diverse materials classes.

Importantly, representation learning has also enabled data efficiency strategies. Transfer learning allows models pretrained on large simulation corpora to be adapted to specialized domains with limited labeled data. Few-shot and meta-learning approaches further enhance predictive performance under data scarcity, leveraging shared structural priors across materials families [12, 22, 28]. These advances collectively position representation learning as both an epistemic compression mechanism and a scalability enabler within materials informatics ecosystems.

Graph neural networks and deep learning architectures in materials contexts

Graph neural networks (GNNs) represent the architectural backbone of contemporary representation learning in materials science. Through iterative message-passing operations, GNNs propagate information across atomic connectivity graphs, enabling the modeling of many-body interactions and emergent structural effects. This relational learning capacity has proven effective in predicting elastic moduli, fracture resistance, diffusion pathways, and thermodynamic stability across diverse materials systems [4, 5].

In polycrystalline contexts, GNNs have been applied to model grain boundary interactions and microstructural heterogeneity, capturing mesoscale effects derived from atomistic simulations. At the electronic scale, crystal graph networks predict bandgaps, formation energies, and density-of-states features directly from simulated structures. Such architectures derive their predictive power from the structural fidelity of simulation outputs that define their training graphs.

Parallel to these developments, physically informed neural networks have emerged to embed governing equations and conservation laws into learning processes. By constraining model outputs through simulation-derived physical relationships, these networks enhance extrapolative robustness and mitigate purely data-driven overfitting [13]. Physics integration may occur through loss-function regularization, architecture design, or hybrid simulation–learning coupling.

Attention-based mechanisms have further refined deep learning architectures in materials contexts. By assigning differential weights to atomic interactions, attention layers enable models to focus on physically salient bonding environments or defect sites, enhancing interpretability and predictive resolution [5]. Such mechanisms are particularly valuable in heterogeneous systems where localized interactions dominate macroscopic properties.

Across these architectures, training data provenance remains consistent: high-throughput simulation infrastructures—primarily DFT and MD—provide the structural and energetic corpora upon which models learn [9]. As a result,

architectural advances remain tightly coupled to the assumptions and approximations embedded within simulation pipelines.

High-throughput computation, autonomous discovery, and closed-loop systems

High-throughput computation has enabled the scaling of simulation from isolated case studies to industrial-scale data generation infrastructures. Automated workflows orchestrate structure generation, relaxation, property calculation, and database curation, producing standardized datasets for machine learning integration [9, 10]. These infrastructures operationalize discovery as a pipeline—one capable of evaluating thousands to millions of candidate materials across compositional and structural spaces.

The convergence of high-throughput simulation with machine learning has catalyzed the emergence of autonomous discovery systems. In these architectures, predictive models guide simulation prioritization, experimental synthesis, and validation cycles, forming iterative discovery loops [6-8]. Robotic platforms execute synthesis and characterization tasks, while machine learning systems interpret outputs and refine subsequent acquisition strategies.

Bayesian active learning frameworks play a central role in this paradigm. By quantifying predictive uncertainty, these methods identify data acquisition targets that maximize information gain. Hierarchical learning approaches further structure acquisition decisions across scales—selecting whether to simulate, experimentally validate, or refine model architectures [8, 10, 21].

Closed-loop infrastructures thus transform materials discovery into an adaptive control system. Error-correction mechanisms reconcile discrepancies between prediction and observation, while iterative retraining refines model fidelity [11, 19, 20]. However, these loops also recirculate simulation priors. When data generation remains simulation-dominant, embedded assumptions propagate through successive iterations, shaping exploration pathways and discovery probabilities.

Inverse materials design and multimodal datasets

Inverse materials design reconfigures predictive modeling into generative exploration. Rather than estimating properties of known structures, inverse frameworks seek to identify structures that satisfy predefined performance criteria. Variational autoencoders, generative adversarial networks, diffusion models, and reinforcement learning architectures enable navigation of high-dimensional design spaces [16-18, 31, 32].

Simulation priors play a critical constraining role in these systems. Generative models are typically trained on simulation-derived datasets, meaning that the latent design spaces they construct reflect the structural and energetic distributions encoded within computational outputs. While such constraints enhance physical plausibility, they also delimit exploration—potentially excluding materials configurations absent from simulation corpora.

Multimodal data integration partially mitigates this limitation. By combining simulation outputs with experimental descriptors, processing conditions, or spectroscopic signatures, models gain exposure to broader physical realities [12]. Fusion strategies align atomistic, microstructural, and thermodynamic modalities within unified embedding spaces, strengthening predictive resilience.

Small-data learning strategies complement multimodal integration. Domain-of-applicability mapping identifies regions where simulation-trained models remain reliable, while transfer learning extends predictive capacity into underrepresented domains [15, 22, 28]. Together, these approaches seek to balance simulation-derived priors with empirical grounding.

Foundation models, simulation–experiment coupling, and epistemic considerations

Recent advances in scientific foundation models extend representation learning to unprecedented scales. Pretrained on massive simulation corpora spanning chemical, structural, and property spaces, these models aim to provide generalizable predictive infrastructures for materials science [23]. Their scale enables cross-domain transfer, few-shot adaptation, and multimodal reasoning.

However, foundation models also aggregate simulation priors at scale. Assumptions embedded across contributing datasets become statistically internalized within pretrained representations. As these models are deployed in downstream discovery workflows, inherited biases may shape predictions in ways that remain opaque without infrastructural interrogation.

Coupling simulation predictions with experimental validation introduces additional epistemic complexity. Discrepancies arise from finite-temperature effects, synthesis imperfections, measurement uncertainty, and real-world processing conditions absent from simulation environments [14]. Bridging these gaps requires calibration frameworks capable of reconciling computational approximations with empirical observations.

Explainable AI and uncertainty quantification provide partial solutions. Feature attribution methods surface influential descriptors, while probabilistic modeling quantifies predictive confidence and epistemic variance [24-26]. Yet these tools operate primarily at the model-output interface, rather than addressing upstream simulation assumptions embedded in data generation itself.

Synthesis and conceptual gap

Across these intersecting literatures—representation learning, graph architectures, high-throughput infrastructures, autonomous systems, inverse design, multimodal fusion, and foundation modeling—a unifying dependency emerges: simulation-generated data serve as the epistemic substrate of materials AI. Machine learning systems, regardless of architectural sophistication, remain conditioned on the physics approximations encoded within their training corpora.

While prior scholarship has advanced interpretability, physical consistency, and experimental integration, the infrastructural role of simulation priors remains conceptually diffuse. Hidden assumptions propagate across representation spaces, influence generative design boundaries, and recirculate through closed-loop discovery systems without systematic articulation.

This synthesis underscores the need for integrative conceptual frameworks that treat simulation priors as dynamic, governable elements within discovery ecosystems. Such frameworks must operate across data generation, representation learning, model inference, and autonomous steering layers—illuminating how physics assumptions shape epistemic trajectories in materials AI.

The Simulation Prior Articulation Framework (SPAF) is positioned to address this gap by embedding simulation priors within a layered, feedback-enabled infrastructural architecture—enabling their identification, modulation, and alignment with broader discovery objectives.

Proposed conceptual framework

The Simulation Prior Articulation Framework (SPAF) is an original conceptual infrastructure designed to systematically address hidden physics assumptions as steerable simulation priors within machine learning materials models. SPAF structures the discovery pipeline into interconnected layers that promote explicit articulation of priors, their integration into representations, and dynamic refinement through feedback.

The framework comprises five structural layers: (1) Multimodal Data Aggregation, which ingests simulation trajectories, high-throughput results, and experimental descriptors; (2) Simulation Prior Extraction and Articulation, where latent assumptions—such as those from energy functionals or boundary conditions—are identified and parameterized; (3) Model Integration, embedding articulated priors into graph neural networks or representation learners; (4) Inference and Discovery Pipeline, generating predictions and guiding inverse design or autonomous selection; and (5) Epistemic Feedback and Steering, which evaluates discrepancies and propagates refinements back to earlier layers.

Data → Model → Discovery pipelines flow unidirectionally through these layers, beginning with aggregated datasets that carry implicit simulation priors, proceeding to models that learn representations conditioned on articulated priors, and culminating in discovery outputs such as candidate structures or experiments. Computational steering logics operate primarily in the feedback layer, using uncertainty estimates and information-theoretic measures to prioritize prior refinement or additional simulation targets. Feedback loops are bidirectional: forward loops propagate refined priors to improve inference robustness, while backward loops adjust data aggregation or extraction parameters based on discovery outcomes. These loops embody workflow dynamics that balance exploration of simulation fidelity against exploitation of available data volumes.

The interaction between articulated priors and data-driven components can be conceptualized as

$$R\left(d, Part\right) = f\left(d;\theta\right) \ominus Part \quad (1)$$

where R denotes the refined representation, f the base learning function, d the input data, θ learned parameters, and ⊖ denotes a modular subtraction operation that explicitly removes or conditions on the articulated prior $Part$. This formalization captures the disentanglement of hidden assumptions from core inference.

The steering dynamics of feedback loops may be expressed as $\Delta P = \eta \cdot \nabla P\left[H\left(Dmodel \mid Dexp\right) + \lambda \cdot Usim\right]$ where ΔP is the prior update, η a steering rate, H the cross-entropy divergence between model and experimental distributions, Usim simulation uncertainty, and λ an infrastructure trade-off parameter. This equation represents the gradient-based refinement that minimizes epistemic misalignment.

Figure 1 depicts SPAF as a layered discovery infrastructure in which simulation-generated data are paired with an explicit prior articulation layer, enabling prior-conditioned representation learning and feedback-driven steering to mitigate epistemic bias propagation.

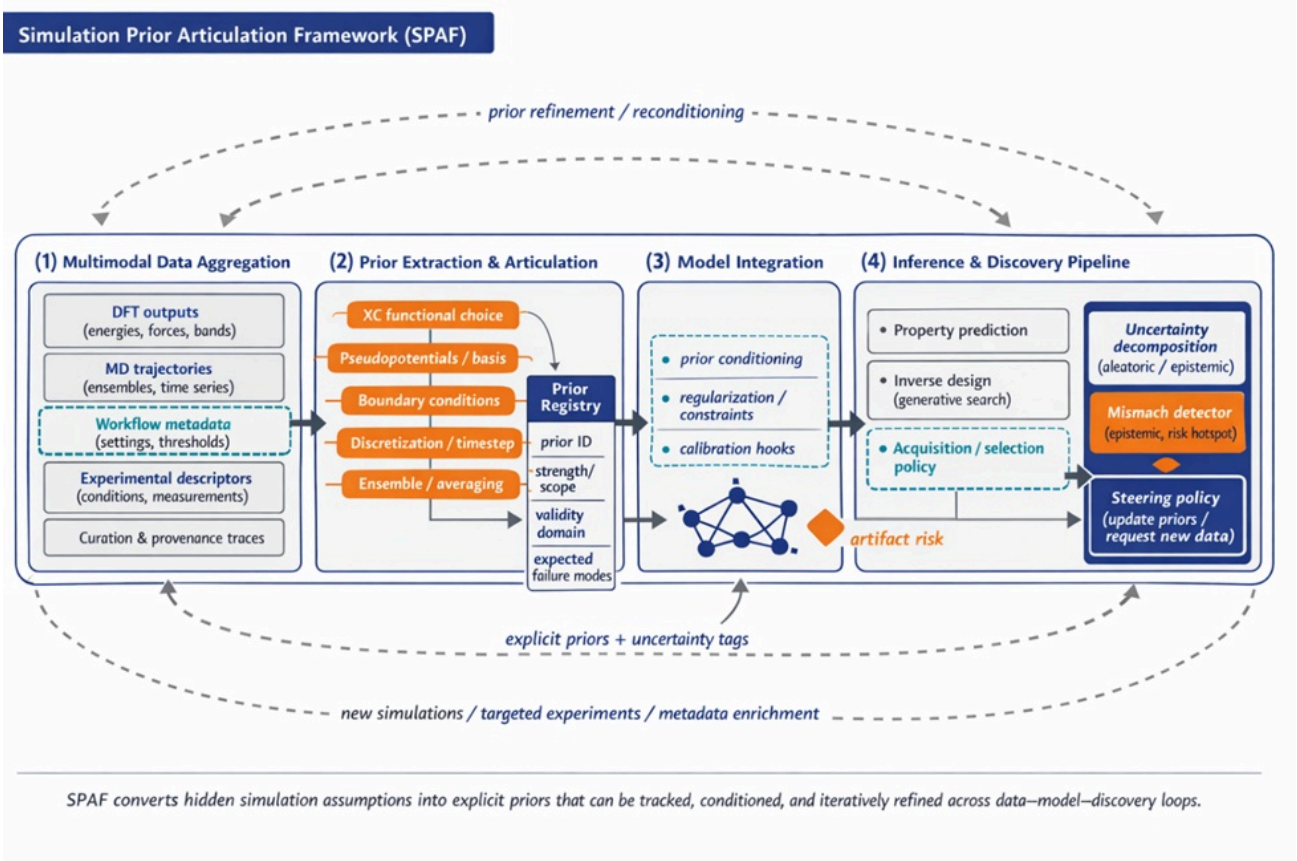

Figure 1. The Simulation Prior Articulation Framework (SPAF) for managing hidden physics assumptions in materials machine learning.

SPAF structures data-driven materials discovery as a five-layer pipeline spanning multimodal data aggregation, explicit extraction and parameterization of simulation priors, integration of articulated priors into representation learning and graph-based models, downstream inference and discovery actions (including inverse design and acquisition policies), and an epistemic steering layer that uses uncertainty and simulation–experiment mismatch signals to refine priors and target new data. Inner feedback loops update priors and model conditioning, while an outer loop drives simulation–experiment coupling by directing targeted computations and measurements, transforming implicit assumptions into traceable, steerable epistemic components of the discovery infrastructure.

Table 2 operationalizes SPAF as an infrastructure logic by mapping each layer to its objective, outputs, and risk-control function.

Table 2. SPAF layer-by-layer design logic: objectives, information products, and epistemic risk controls

SPAF layer	Primary objective	Key inputs	Output "information products"	Epistemic risk controlled	Steering signal that triggers refinement
(1) Multimodal Data Aggregation	Consolidate simulation + experiment with provenance	DFT/MD/phase-field outputs; experimental descriptors; workflow metadata	Curated datasets + provenance traces; modality alignment	Hidden assumptions buried in provenance gaps	Missing metadata; modality inconsistency; coverage gaps

(2) Prior Extraction & Articulation	Convert implicit assumptions into explicit priors	Computational settings; approximations; curation rules	Prior registry (IDs, scope, strength, validity domain, failure modes)	Untracked simulation bias and assumption drift	Repeated mismatch patterns; unstable uncertainty under shift
(3) Model Integration	Condition representation learning on articulated priors	Priors + training data + representation architecture	Prior-aware embeddings; calibrated constraints; modular prior hooks	Artifact propagation through embeddings	Embedding drift; sensitivity spikes; OOD indicators
(4) Inference & Discovery Pipeline	Produce actionable predictions and designs	Prior-aware model outputs; objectives/constraints	Ranked candidates; inverse designs; acquisition policies	Overconfident steering; spurious optimization	High epistemic uncertainty; divergence in candidate outcomes
(5) Epistemic Feedback & Steering	Detect mismatch and update priors/data targets	Prediction–observation comparisons; uncertainty decomposition	Prior updates; targeted simulation/experiment requests	Epistemic misalignment (sim↔exp) and loop reinforcement	Cross-entropy/divergence rise; persistent residual structure

Through these elements, SPAF provides interpretive insights into how hidden physics assumptions can be transformed from opaque biases into transparent, computationally steerable components of the materials discovery infrastructure.

Analytical implications

Representation–Inference interactions under SPAF

Within the Simulation Prior Articulation Framework (SPAF), representation learning gains enhanced robustness through the explicit handling of simulation priors, altering how graph neural networks and deep architectures process materials data. By disentangling priors from raw simulation outputs, representations become less susceptible to propagation of hidden assumptions, such as those from periodic boundary conditions or exchange-correlation functionals. This interaction fosters a more modular inference process, where predictions for properties like stability or mechanical behavior can incorporate adjustable prior strengths, leading to tunable epistemic confidence in downstream applications.

In inverse design contexts, SPAF's layered structure implies that articulated priors guide generative processes by constraining latent spaces derived from multimodal datasets. For instance, when steering toward novel structures, the framework's feedback mechanisms allow for iterative refinement of representations, balancing simulation-derived constraints against experimental discrepancies. This dynamic underscores computational workflow efficiencies, where prior articulation reduces the need for exhaustive retraining, instead favoring targeted updates to model components.

Discovery steering logics and feedback dynamics

SPAF introduces steering logics that operate at the intersection of data aggregation and inference layers, enabling adaptive control over discovery pipelines. Feedback loops, informed by divergence measures between simulation and experimental distributions, facilitate real-time adjustments to prior parameterizations. This implies a shift in autonomous systems toward infrastructures that prioritize epistemic alignment, potentially optimizing closed-loop cycles by minimizing iterations spent on misaligned assumptions.

The trade-offs in computational resources become evident here: higher fidelity in prior extraction demands increased processing in early layers, yet yields streamlined discovery in later stages. For high-throughput setups, this logic suggests

integrating SPAF to filter simulation outputs proactively, enhancing the utility of graph-based models in identifying viable candidates without amplifying latent biases.

The epistemic risk structure under SPAF can be conceptualized as

$$E = \sum iwi \cdot (\overline{D}(Part, i \mid \ \ Ptrue \mid) + \beta \cdot Ccomp, i) \quad (2)$$

where E is the overall epistemic risk, weights for each prior component, D a divergence metric between articulated and true priors, β a balancing factor, and Ccomp,i the computational cost of articulation. This expression captures the interaction between accuracy in prior handling and infrastructure demands, highlighting pathways for risk mitigation.

Infrastructure trade-offs in simulation–experiment coupling

Applying SPAF to coupled simulation–experiment ecosystems reveals trade-offs between data volume from high-throughput computations and the depth of prior articulation. In foundation models pretrained on vast simulation corpora, unaddressed priors may inflate epistemic risks during fine-tuning on experimental data. SPAF's framework implies a reconfiguration of these models to include dedicated prior layers, fostering better transferability across domains.

For materials informatics pipelines, this leads to interpretive insights on how articulated priors can bridge gaps in small-data regimes, where simulation assumptions often dominate. The steering logics further imply enhanced resilience in active learning scenarios, allowing systems to dynamically allocate resources toward resolving high-risk priors, thus optimizing overall discovery efficiency without empirical dependencies.

Results and Discussion

The SPAF framework, by centering simulation priors as explicit elements, offers a pathway to more integrated computational materials engineering. It aligns with ongoing efforts in explainable machine learning [24-26] and physically informed architectures [13, 14], yet extends these by embedding priors within a systems-level infrastructure. This approach mitigates risks in representation learning [3, 29, 30] and inverse design [16-18, 31, 32], where hidden assumptions could otherwise undermine reliability.

In autonomous discovery and closed-loop systems [6-8, 11, 19-21], SPAF's feedback dynamics provide tools for epistemic steering, complementing Bayesian and hierarchical methods [8, 10, 21]. The framework's emphasis on workflow dynamics addresses challenges in multimodal data handling [11, 12] and domain applicability [15], promoting infrastructures that adapt to evolving simulation–experiment interfaces.

While SPAF remains conceptual, its implications suggest avenues for enhancing foundation models [23] and small-data strategies [22, 28], ensuring that data-driven paradigms in materials science [1, 2] evolve toward greater transparency and robustness. Trade-offs identified underscore the need for balanced computational investments, aligning with community perspectives on machine learning adoption [27].

Conclusion

The Simulation Prior Articulation Framework (SPAF) provides a novel conceptual infrastructure for managing hidden physics assumptions in machine learning materials models. By delineating layered pipelines, feedback loops, and steering logics, SPAF offers systems-level insights into representation–inference interactions and epistemic risk structures. This advances computational materials engineering toward more transparent discovery ecosystems, with

implications for materials informatics, inverse design, and autonomous systems. Future integrations could further refine data-driven workflows, emphasizing the transformative role of articulated simulation priors.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 16 Jan 2023 Revised: 23 Mar 2023 Accepted: 01 May 2023
Published online: 18 September 2023

References

Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. npj Comput Mater. 2019;5:83.

Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. npj Comput Mater. 2017;3:54.

Fung V, Zhang J, Juarez E, Sumpter BG. Benchmarking graph neural networks for materials chemistry. npj Comput Mater. 2021;7(1):84.

Kojic M, Zecevic M, Cukovic M, Grujovic N, Reese SP, Knezevic M. Graph neural networks for efficient learning of mechanical properties of polycrystals. Comput Mater Sci. 2023;217:111894.

Schmidt J, Shi H-L, Isnard O, Marques MAL, Botti S. Crystal graph attention networks for the prediction of stable materials. Sci Adv. 2021;7(49):eabi7948.

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.

Szymanski NJ, Rendy B, Fei Y, Kumar RE, He T, Milsted D, McDermott MJ, et al. An autonomous laboratory for the accelerated synthesis of inorganic materials. Nature. 2023;624(7990):86-91.

Kusne AG, Yu H, Wu C, Zhang H, Hattrick-Simpers J, DeCost B, et al. On-the-fly closed-loop materials discovery via Bayesian active learning. Nat Commun. 2020;11(1):5966

Wolloch M, Losi G, Chehaimi O, Yalcin F, Ferrario M, Righi MC. High-throughput generation of potential energy surfaces for solid interfaces. Comput Mater Sci. 2022;207:111302.

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.

Kalantre S, Ray S, Mishra AA, Ndayishimiye A, Bud'ko SL, Canfield PC, et al. Closed-loop superconducting materials discovery. npj Comput Mater. 2023;9(1):156.

Gupta V, Choudhary K, Tavazza F, Campbell C, Liao W-k, Choudhary A, et al. Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data. Nat Commun. 2021;12(1):6595.

Pun GPP, Yamakov V, Mishin Y. Physically informed artificial neural networks for atomistic modeling of materials. Nat Commun. 2019;10(1):2339.

Yaghoobi M, Adnan A, Hartmaier A. Analyses of internal structures and defects in materials using physics-informed neural networks. Sci Adv. 2022;8(7):eabk0644.

Sutton C, Boley M, Ghiringhelli LM, Rupp M, Vreeken J, Scheffler M. Identifying domains of applicability of machine learning models for materials science. Nat Commun. 2020;11(1):4428.

Jha D, Gupta V, Ward L, Yang Z, Wolverton C, Foster I, et al. Attribute driven inverse materials design using deep learning Bayesian framework. npj Comput Mater. 2019;5(1):127.

Attari V, Khatamsaz D, Allaire D, Arroyave R. Towards inverse microstructure-centered materials design using generative phase-field modeling and deep variational autoencoders. Acta Mater. 2023;259:119204.

Deng C, Liu C, Hu J, Huang K, Zhao Y, Ye W, et al. Rapid inverse design of metamaterials based on prescribed mechanical behavior through machine learning. Nat Commun. 2023;14(1):5565.

Kavalsky L, Hegde VI, Hummelshøj JS, Johnson MS, Meredig B, Viswanathan V. By how much can closed-loop frameworks accelerate computational materials discovery?. Digit Discov. 2023;2(4):1168-77.

Choubisa H, Berman A, Kenis PJA, Ertekin E. Closed-Loop error-correction learning accelerates experimental discovery of thermoelectric materials. Adv Mater. 2023;35(40):2302575.

Ament S, Amsler M, Sutherland DR, Chang M-C, Guevarra D, Connolly AB, et al. Autonomous materials synthesis via hierarchical active learning of nonequilibrium phase diagrams. Sci Adv. 2021;7(51):eabg4930.

Zhao Y, Krishnan NMA, Bauchy M, Lu W. Small data machine learning in materials science. npj Comput Mater. 2023;9(1):42.

Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R, et al. Recent advances and applications of deep learning methods in materials science. npj Comput Mater. 2022;8(59).

Kailkhura B, Gallagher B, Kim S, Hiszpanski A, Han TY-J. Reliable and explainable machine-learning methods for accelerated material discovery. npj Comput Mater. 2019;5(1):108.

Zhong X, Gallagher B, Liu S, Kailkhura B, Hiszpanski A, Han TY-J. Explainable machine learning in materials science. npj Comput Mater. 2022;8(1):204.

Pilania G. Machine learning in materials science: From explainable predictions to autonomous design. Comput Mater Sci. 2021;193:110360.

Boyce B, Dingreville R, Desai S, Walker E, Shilt T, Bassett KL, et al. Machine learning for materials science: Barriers to broader adoption. Matter. 2023;6(5):1320-3.

Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. npj Comput Mater. 2018;4(1):25.

Goodall REA, Lee AA. Predicting materials properties without crystal structure: deep representation learning from stoichiometry. Nat Commun. 2020;11(1):6280.

Fung V, Hu G, Ganesh P, Sumpter BG. Distributed representations of atoms and materials for machine learning. npj Comput Mater. 2022;8(1):40.

Cubuk ED, Schoenholz SS, Rieser JM, Malone B, Rottler J, Durian DJ, et al. Inverse design of solid-state materials via a continuous representation. Matter. 2019;1(5):1370-84.

Coli GM, Baldi E, de Oliveira LFCF, Dijkstra M. Inverse design of soft materials via a deep learning–based evolutionary strategy. Sci Adv. 2022;8(3):eabj6731.