ORIGINAL RESEARCH Open access

Transfer Learning Across Materials Classes: Conceptual Boundaries of Reusability

Lucas Meyer[1*], Stefan Braun[1], Anna Schmid[2]

Abstract

In the evolving landscape of computational and data-driven materials engineering, transfer learning has emerged as a pivotal strategy to address data scarcity and enhance predictive capabilities across diverse materials systems. This approach leverages pre-trained models from one materials class to inform modeling in another, capitalizing on shared representational structures within high-dimensional chemical and physical spaces. However, the conceptual boundaries of reusability remain underexplored, particularly in terms of how representational invariances and domain shifts influence cross-class applicability. This manuscript introduces a novel conceptual framework, termed the Reusability Boundary Architecture (RBA), which delineates the systemic interactions between data representations, model architectures, and discovery workflows in transfer learning paradigms. By integrating insights from materials informatics, graph neural networks, and uncertainty quantification, the RBA elucidates the epistemic trade-offs inherent in transferring knowledge across materials classes, such as from inorganic crystals to organic polymers or metallic alloys to ceramics. The framework emphasizes computational steering logics that dynamically adjust for feature misalignment and contextual divergences, fostering more robust integration of simulation and experimental pipelines. Implications for the field include enhanced design of multimodal datasets, refined autonomous discovery systems, and improved inverse materials engineering, ultimately accelerating innovation in sustainable materials development without relying on empirical validations. This work provides a theoretical foundation for navigating the reusability frontiers in computational materials science, promoting interdisciplinary synergies between machine learning and domain-specific knowledge.

Keywords Data-driven discovery, Transfer learning, Representation learning, Domain adaptation, Computational materials science, Materials reusability

*Correspondence:
Lucas Meyer
lucas.meyer@gmail.com
[1] Department of Materials Modeling and Simulation, Faculty of Engineering, ETH Zurich, Zurich, Switzerland
[2] Department of Data-Driven Materials Science, Faculty of Engineering, University of Bern, Bern, Switzerland

Introduction

The advent of data-driven paradigms in materials engineering

The integration of computational methodologies with data-driven paradigms has catalyzed a structural transformation in materials engineering, reshaping discovery logics from empirically guided experimentation toward predictive, informatics-driven design. Historically, materials innovation progressed through iterative trial-and-error experimentation, where compositional adjustments and process refinements were validated through labor-intensive characterization cycles. While foundational breakthroughs emerged from this paradigm, its epistemic velocity was constrained by slow experimental throughput, fragmented data ecosystems, and limited predictive foresight. The contemporary convergence of high-performance computing, automated experimentation, and machine learning has reconfigured this landscape, enabling materials discovery to operate within statistically grounded inference infrastructures rather than isolated empirical observations.

This paradigm leverages expansive datasets generated through high-throughput computational screening, combinatorial synthesis platforms, and robotic characterization pipelines to uncover latent patterns governing structure–property relationships [1, 2]. Automated density functional theory (DFT) workflows, for instance, now produce large-scale repositories of formation energies, electronic structures, and thermodynamic descriptors, while high-resolution imaging and spectroscopy platforms capture microstructural and functional signatures across scales. When integrated within materials informatics pipelines, these multimodal datasets form the substrate upon which machine learning systems extract predictive correlations. Such infrastructures enable accelerated identification of candidate materials for applications spanning semiconductors, catalysis, structural alloys, energy storage systems, and biomaterials.

Central to this evolution is the application of machine learning techniques capable of modeling nonlinear and multiscale interactions embedded within materials systems. Supervised learning architectures infer mappings between structural descriptors and functional outputs, while unsupervised approaches reveal clustering patterns that signal phase boundaries or compositional affinities. More recently, generative and inverse design frameworks have begun navigating latent chemical design spaces to propose novel materials configurations. These developments collectively shift materials engineering toward a predictive science—one where discovery is guided not solely by experimental iteration but by algorithmic foresight grounded in statistical inference.

However, the heterogeneity of materials classes introduces profound modeling challenges. Materials systems encompass crystalline solids, amorphous glasses, polymers, composites, nanomaterials, and bio-hybrid constructs—each governed by distinct bonding regimes, dimensional constraints, and environmental sensitivities. Crystalline materials exhibit periodic symmetry and long-range order; polymers manifest chain entanglement and conformational flexibility; biomaterials operate within solvent-mediated, temperature-responsive contexts. Such diversity complicates the generalization of predictive models across domains, as representations optimized for one class may inadequately capture the physicochemical dynamics of another. Consequently, the scalability of data-driven discovery is bounded by representational fragmentation across materials subclasses.

Transfer learning, a specialized subset of machine learning, emerges as a computational strategy to mitigate this fragmentation. Rather than training predictive systems de novo for each materials domain, transfer learning repurposes knowledge encoded within pretrained models to inform new tasks. Representational features learned in data-rich environments—such as large crystalline databases—can be fine-tuned to operate in data-sparse domains, reducing the need for exhaustive data generation [3-5]. This paradigm is particularly consequential in materials science, where experimental validation remains resource-intensive and rare materials phenomena lack extensive labeled datasets. By initializing models with pretrained weights, transfer learning accelerates convergence, enhances predictive stability, and expands discovery reach into underexplored chemical territories.

Despite these advantages, the reusability of transferred knowledge is bounded by conceptual and physicochemical limits that govern representational compatibility. Knowledge embedded within a trained model reflects the statistical distributions, bonding topologies, and environmental assumptions of its source domain. When transferred into divergent materials contexts, these embedded priors may lose interpretive validity. For instance, atomic-scale descriptors effective in metallic systems—where delocalized electron behavior dominates—may fail to generalize to organic or polymeric materials characterized by localized covalent bonding and conformational variability [6, 7]. Similarly, representations trained on bulk crystalline datasets may inadequately describe surface-dominated nanostructures or defect-rich amorphous phases. These discontinuities underscore the existence of implicit reusability boundaries that regulate when and how transfer learning can be effectively operationalized across materials classes.

Understanding these boundaries requires a systems-level interrogation of representational alignment. It is insufficient to evaluate transfer learning solely through performance metrics; one must examine descriptor congruence, embedding topology overlap, and physicochemical invariance preservation across domains. Without such alignment, transferred models risk introducing interpretive distortions rather than predictive insights. Consequently, delineating the structural conditions under which transfer learning enhances discovery—and those under which it induces epistemic risk—emerges as a foundational research imperative. paving the way for more adaptive computational workflows in materials engineering **(Figure 1)**.

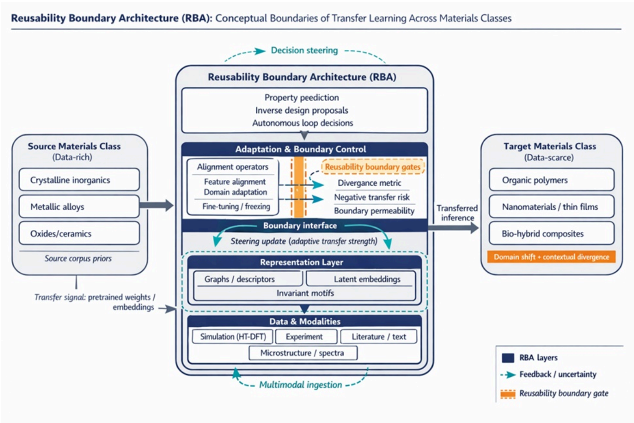

Figure 1. Reusability Boundary Architecture (RBA): A Conceptual Framework for Transfer Learning Across Materials Classes

Challenges in cross-class knowledge transfer

Data scarcity persists as one of the most significant structural bottlenecks in contemporary materials science. While computational infrastructures generate high-volume datasets for well-studied materials families, experimental validation remains costly, time-intensive, and instrumentation-dependent [8, 9]. Emerging materials classes—such as bio-inspired polymers, quantum materials, or extreme-environment ceramics—often lack the data density required for robust standalone model development. Transfer learning offers a pathway to mitigate this imbalance by initializing predictive architectures with pretrained weights derived from related domains, thereby enhancing convergence rates and predictive fidelity under limited data conditions [10-12].

However, the efficacy of cross-class transfer is contingent upon the alignment of feature spaces between source and target domains. When descriptor distributions diverge substantially, transferred knowledge may encode irrelevant correlations or suppress domain-specific signals. This phenomenon—commonly termed negative transfer—can degrade predictive performance relative to models trained exclusively on target data [13, 14]. For example, transferring representations trained on oxide ceramics to metallic glass systems may obscure disorder-driven structural effects critical to accurate property prediction. Thus, transfer learning must be conceptualized not as universally beneficial but as conditionally effective within bounded representational regimes.

Literature documents selective successes in cross-class adaptation, including the transfer of models from bulk crystalline systems to thin films, or from oxide chemistries to multicomponent alloys [15-17]. Yet such implementations often operationalize transfer as a pragmatic shortcut rather than interrogating its systemic limits. The absence of formal frameworks delineating reusability boundaries restricts interpretability, reproducibility, and scalability across discovery infrastructures.

The integration of multimodal data further complicates transfer processes. Contemporary materials datasets increasingly combine spectroscopic measurements, microstructural imaging, thermodynamic simulations, and textual synthesis metadata. Each modality introduces distinct noise structures, resolution scales, and epistemic uncertainties. When transfer learning operates across such heterogeneous inputs, uncertainty propagation can distort inference reliability [18-20]. Representations pretrained on crystallographic graphs, for instance, may not seamlessly integrate spectroscopic embeddings or morphological descriptors. Accordingly, uncertainty quantification becomes indispensable, enabling predictive systems to calibrate confidence in relation to domain divergence and data fidelity [21, 22].

These concerns are amplified within autonomous discovery ecosystems. Closed-loop platforms—where AI systems propose candidates, robotic platforms synthesize them, and characterization outputs feed back into retraining cycles—depend on reliable cross-domain inference [23, 24]. If transferred models misrepresent uncertainty or encode incompatible priors, discovery loops may converge toward experimentally infeasible or suboptimal design regions. Consequently, the structural limits of transfer learning reusability directly shape the epistemic reliability of autonomous materials innovation systems.

Emerging computational architectures and their limitations

Recent advances in deep learning architectures have expanded the operational scope of transfer learning across materials classes. Graph neural networks (GNNs) have become particularly influential due to their capacity to encode atomic connectivity, coordination environments, and symmetry invariances within relational graph structures [25-27]. By representing materials as node-edge systems, GNNs preserve rotational symmetry, permutation invariance, and local bonding topology—features critical for transferable representation learning.

Latent embeddings generated through such architectures capture hierarchical physicochemical abstractions, enabling fine-tuned adaptation across related materials systems [28, 29]. Multi-fidelity learning frameworks further augment transfer potential by integrating high-volume, low-accuracy simulation data with sparse, high-accuracy experimental datasets. These hybrid learning environments enhance predictive robustness while reducing dependence on costly data acquisition pipelines.

Yet the conceptual boundaries of reusability extend beyond architectural design. Transfer efficacy is mediated by data curation practices, training corpus composition, and inference workflow integration. Foundation models trained on large-scale scientific datasets illustrate this tension. While they encode broad chemical knowledge, their training distributions often overrepresent stable crystalline materials and underrepresent disordered, metastable, or synthesis-sensitive systems [30-32]. Consequently, their cross-class transferability may be uneven, performing reliably within familiar compositional regimes while faltering in emergent or underrepresented materials domains.

High-throughput computational infrastructures may inadvertently reinforce these biases. Datasets generated for targeted discovery campaigns—such as catalytic screening or battery electrode optimization—embed selection criteria that shape learned representations [3, 4]. When such representations are transferred into unrelated design contexts, embedded biases may constrain exploration rather than enable it. This issue becomes particularly consequential in inverse materials design, where predictive systems must extrapolate beyond observed chemical spaces to engineer materials from desired functional properties [11]. Accordingly, these constraints can be organized into a boundary taxonomy that clarifies where transfer succeeds, where it fails, and why negative transfer emerges across heterogeneous materials classes **(Table 1)**.

Table 1. Taxonomy of reusability boundaries in transfer learning across materials classes

Reusability boundary type	What diverges across classes	Typical symptom in transfer
Representational mismatch	Descriptors encode incompatible physics (e.g., periodic crystals vs conformational polymers)	Apparent accuracy on source, collapse on target
Invariant violation	Symmetries/topologies assumed by model not valid in target	Instability under small perturbations
Distributional domain shift	Feature distributions differ (composition, defects, morphology)	Negative transfer; degraded calibration
Modal mismatch	Source largely simulation; target largely experimental (noise, bias)	Overconfident predictions
Scale discontinuity	Atomic-scale source vs meso-/microstructural target	Missing mechanisms; spurious correlations
Task misalignment	Source task differs (formation energy) vs target task (processability, durability)	Transfer helps training loss but harms decision utility
Data lineage and curation bias	Source corpus selection bias embeds priors	Systematic omission of viable region

Feedback instability (closed-loop)	Iterative updating amplifies early errors	Premature convergence

Without clearly articulated reusability boundaries, such extrapolative efforts risk inefficiency and epistemic drift. Transferred models may propose chemically implausible compositions or overlook viable design pathways due to representational misalignment. Thus, delineating the structural limits of transfer learning becomes essential for sustaining discovery fidelity across expanding materials design frontiers.

Positioning the conceptual contribution

In response to these systemic gaps, this manuscript advances a conceptual framework that interrogates the structural boundaries governing transfer learning reusability across heterogeneous materials classes. Rather than treating transferability as an empirical byproduct of model training, we conceptualize it as a boundary-regulated systems phenomenon shaped by representational congruence, physicochemical compatibility, and epistemic feedback dynamics.

We introduce the Reusability Boundary Architecture (RBA)—a novel interpretive construct that maps the layered infrastructures through which knowledge transfer unfolds. RBA delineates how descriptor alignment, latent embedding overlap, uncertainty modulation, and domain divergence gradients collectively regulate transfer efficacy. By embedding these processes within closed-loop discovery ecosystems, the framework elucidates how experimental validation feedback, multimodal integration, and autonomous inference reshape reusability thresholds over iterative design cycles.

Through this systems-level lens, transfer learning is reframed not merely as a computational optimization strategy but as a foundational epistemic infrastructure within data-driven materials engineering. Understanding where knowledge travels—and where it fractures—becomes indispensable for designing adaptive, uncertainty-aware discovery pipelines capable of navigating the heterogeneity of modern materials innovation landscapes.

Theoretical Background & Literature Synthesis

Foundational concepts in materials informatics and machine learning

Materials informatics serves as the backbone for data-driven approaches, employing statistical and machine learning methods to analyze and predict materials behaviors from large datasets [1, 2]. Key to this is representation learning, where materials are encoded into machine-readable formats, such as atomic fingerprints or graph-based structures, enabling the extraction of meaningful features [4, 29]. Transfer learning builds upon this by allowing models to adapt pre-learned representations to new tasks, particularly useful in scenarios with limited data [3, 5, 6]. In computational materials science, this often involves fine-tuning models trained on abundant datasets, like those from crystal structure databases, for application in niche areas such as nanomaterials or biomaterials [8, 10, 11].

The reusability of these representations across materials classes depends on shared underlying physics, yet differences in scale—from atomic to mesoscopic—introduce boundaries that limit direct applicability [7, 26]. For instance, graph neural networks excel in capturing local environments in crystalline materials but may require adjustments for amorphous or polymeric systems where long-range interactions dominate [12, 23, 25].

Domain shifts and adaptation strategies in cross-class transfer

Domain shifts represent a core challenge in transfer learning, manifesting as discrepancies in data distributions between source and target materials classes [9, 13, 14]. Literature synthesizes various adaptation techniques, including feature alignment and adversarial training, to mitigate these shifts [16, 18, 30]. In materials contexts, this might involve aligning latent spaces from oxide perovskites to spinel structures, ensuring that transferred features retain predictive power [7, 15, 21].

However, conceptual boundaries emerge when shifts involve not just statistical variances but epistemic differences, such as varying levels of uncertainty in experimental versus simulated data [19, 20, 22]. High-throughput computation generates precise but idealized datasets, while experimental multimodal sources introduce real-world variabilities that complicate transfer [17, 24, 31]. Synthesis of these works reveals a need for systemic views that consider how adaptation strategies interact with discovery pipelines, from data ingestion to model deployment.

Integration of uncertainty and feedback in discovery workflows

Uncertainty quantification plays a pivotal role in bounding reusability, providing measures of confidence in transferred predictions [1, 8, 19]. In data-driven frameworks, this involves propagating uncertainties through neural architectures, ensuring that cross-class transfers do not amplify errors [22, 27]. Autonomous discovery systems exemplify this, where closed-loop mechanisms use uncertainty estimates to steer experimental iterations [23, 24, 32].

Literature underscores the importance of feedback loops in refining transfers, such as iteratively updating models with target-domain data to close reusability gaps [10, 11, 28]. Yet, across materials classes, these loops must account for class-specific dynamics, like phase transitions in alloys versus conformational changes in polymers [12, 16, 25]. Synthesis highlights that while individual studies address tactical adaptations, a broader conceptual integration is lacking, particularly in how uncertainty interacts with representational reusability.

Representation learning and architectural innovations

Advanced architectures, including atom-set models and crystal graph networks, facilitate transfer by learning hierarchical representations that abstract common motifs across classes [4, 25, 26, 29]. These enable reusability through modular designs, where lower layers capture universal features like bond angles, and higher layers adapt to class-specific properties [3, 5, 6, 30].

However, boundaries arise in multimodal contexts, where integrating diverse data types—e.g., spectra with structures—requires sophisticated fusion strategies [17, 18, 30]. Works on foundation models for science suggest scalable reusability, but their application across materials classes reveals trade-offs in specificity versus generality [2, 13, 31]. Synthesis indicates that while architectural innovations expand reusability, they often overlook the epistemic structures that define transfer limits, such as contextual embeddings that vary by class.

Simulation-experiment coupling and inverse design implications

Coupling simulations with experiments enhances transfer learning by bridging idealized computational domains with empirical realities [11, 15, 24]. In inverse design, this coupling allows backward mapping from properties to structures, leveraging transferred knowledge to explore design spaces [7, 9, 28]. Yet, reusability boundaries are evident when simulation-derived models fail to capture experimental nuances, necessitating hybrid workflows [14, 16, 21].

Literature synthesizes that closed-loop systems amplify these implications, using transferred models to prioritize candidates in high-throughput screening [1, 23, 32]. However, without delineating boundaries, such as those imposed by data modality mismatches, transfers risk inefficiency [8, 19, 22]. Overall, the synthesis points to an integrative need for frameworks that conceptualize reusability not as isolated techniques but as interconnected systems influencing materials discovery.

Proposed conceptual framework

The Reusability Boundary Architecture (RBA)

To address the conceptual boundaries of reusability in transfer learning across materials classes, this manuscript proposes the Reusability Boundary Architecture (RBA), an original systemic framework that maps the interactions between data representations, model adaptations, and discovery workflows. The RBA conceptualizes transfer learning as a layered infrastructure, where reusability is governed by dynamic boundaries that emerge from domain alignments and divergences. At its core, the framework structures the process into three primary layers: the Representation Layer, which encodes materials features; the Adaptation Layer, which handles cross-class mappings; and the Discovery Layer, which integrates outputs into iterative pipelines.

In the Representation Layer, data from source materials classes—such as crystalline inorganics—are transformed into latent embeddings that capture invariant properties like atomic coordination. These embeddings feed into the

https://doi.org//r828506645

Adaptation Layer, where boundary functions assess reusability by quantifying shifts in feature distributions. Feedback loops within this layer allow for iterative refinement, adjusting for class-specific contexts like electronic structures in metals versus steric effects in organics. The Discovery Layer then steers computational workflows, using bounded reusability to guide inverse design or autonomous experimentation, ensuring that transferred knowledge enhances rather than hinders epistemic reliability.

Central to the RBA are computational steering logics that dynamically modulate transfer processes. These logics operate through feedback mechanisms, where discrepancies in uncertainty profiles trigger recalibrations, fostering robust cross-class applicability. For instance, in transitioning from alloy datasets to ceramic predictions, the framework's logics prioritize features with high reusability scores, mitigating negative transfer risks.

Data → Model → Discovery Pipelines and Feedback Loops

The RBA delineates a pipeline from data ingestion to discovery outcomes, emphasizing how reusability boundaries shape each stage. Data flows begin with multimodal inputs, processed through representation learning to form transferable embeddings. Model components then apply adaptation strategies, bounded by conceptual thresholds that define reusability frontiers—such as alignment metrics that decay with increasing domain divergence.

Feedback loops are integral, cycling insights from discovery back to representation refinement. This closed-loop dynamic ensures that epistemic risks, like overgeneralization across classes, are managed through iterative boundary assessments. Computational steering logics embedded in these loops use symbolic interactions to balance trade-offs, such as generality versus specificity in model architectures.

One such dynamic can be conceptualized as the reusability trade-off function, expressed as $R = f(D_s, D_t, A)$ where D_s and D_t represent source and target domain distributions, and A captures adaptation strength. This function highlights how reusability R diminishes as domain divergence increases, unless mitigated by A, illustrating the interaction between data alignments and adaptive mechanisms.

Another key interaction is captured in the feedback efficiency metric, which may be expressed as $E = \sum i w_i \cdot (U_i - 1 \cdot F_i)$, with U_i denoting uncertainty in layer i, F_i the feedback contribution, and w_i weights reflecting layer priorities. This formalizes how feedback loops enhance efficiency by inversely scaling with uncertainty, steering pipelines toward reliable cross-class transfers. To make the framework actionable as a systems design logic, the RBA can be expressed as a layer-wise blueprint linking inputs, boundary signals, steering actions, and discovery outputs (Table 2).

Table 2. RBA design blueprint: layer-wise components, signals, and outputs for boundary-aware transfer

RBA layer	**Primary function**	**Inputs**	
Layer 1: Data & Modalities	Assemble cross-domain evidence substrate	Simulation, experiment, literature/text, microstructure/spectra	i nc p
Layer 2: Representation Layer	Construct transferable encodings	Graphs/descriptors; embeddings; invariance constraints	E va mi
Layer 3: Adaptation & Boundary Control	Regulate transfer under divergence	Pretrained weights; target few-shot data; embeddings	D me tı cal
Layer 4: Discovery & Deployment	Convert inference into decisions	Adapted predictor; uncertainty estimates; constraints	ove
Feedback loops (cross-	Stabilize transfer	New target measurements;	U

layer)	over time	errors; constraint violations	dis

A third conceptualization addresses boundary propagation, formulated as $Bp = \prod\_j (1 - \delta_j)$, where δ_j is the divergence factor at junction j in the pipeline. This product captures the cumulative effect of boundaries, showing how small divergences amplify through stages, necessitating steering logics to maintain overall reusability.

These formulas underscore the RBA's interpretive power, framing reusability not as static but as emergent from systemic interactions.

Analytical implications

The Reusability Boundary Architecture (RBA) offers interpretive lenses for dissecting the systemic implications of transfer learning across materials classes, revealing how boundary dynamics influence computational workflows. Analytically, the framework underscores the interplay between representation alignments and epistemic risks, where reusability boundaries act as filters that modulate knowledge flow. In materials informatics pipelines, this implies that cross-class transfers must navigate trade-offs between feature generality and domain specificity, potentially reshaping data curation strategies to prioritize invariant motifs that enhance boundary permeability.

For instance, in high-throughput computation scenarios, the RBA's layers suggest that source-domain embeddings from crystalline materials can inform polymeric systems only if adaptation logics account for scalar divergences, such as differing length scales. This analytical perspective highlights infrastructure trade-offs: overly rigid boundaries may stifle innovation in inverse design, while porous ones risk introducing noise that erodes predictive fidelity [3, 4, 6]. By framing these as dynamic interactions, the RBA steers toward hybrid workflows where simulation-experiment coupling benefits from bounded reusability, ensuring that transferred models contribute to closed-loop optimization without amplifying uncertainties [1, 11, 23].

Furthermore, the framework's feedback loops imply enhanced discovery steering, where epistemic risk structures guide the selection of transferable features. In representation learning contexts, this means interpreting graph neural network outputs through boundary assessments, fostering more adaptive architectures that evolve with multimodal inputs [12, 25, 29]. Analytically, this interaction can be expressed as the alignment efficiency dynamic, conceptualize $A_e = \int g(R_d), dM$, where R_d denotes representational divergence and M the model manifold, capturing how efficiency integrates over adaptive paths to minimize cross-class mismatches.

Another implication arises in uncertainty quantification, where the RBA posits that boundary propagation affects inference reliability across classes like alloys to perovskites [7, 19, 22]. This suggests computational logics that dynamically weight feedback based on risk profiles, optimizing for scenarios where data scarcity in target domains is offset by source richness [5, 8, 10]. Such insights extend to autonomous discovery systems, implying that reusability boundaries can be leveraged to prioritize exploration in uncharted materials spaces, balancing exploration-exploitation in design pipelines [15, 16, 25].

Overall, these analytical implications position the RBA as a tool for interpreting representation-inference interactions, where trade-offs in reusability inform the design of foundation models and multimodal datasets [13, 18, 30, 31]. By emphasizing systemic insights over isolated tactics, the framework encourages computational materials engineering to adopt boundary-aware strategies, ultimately refining the epistemic foundations of data-driven innovation.

Results and Discussion

The conceptual boundaries delineated by the RBA invite a broader discourse on the reusability paradigms in computational and data-driven materials engineering. Integrating literature syntheses, the framework reveals that while transfer learning has advanced predictive analytics across classes [3, 5, 12], systemic oversights in boundary management often limit its scope. For example, adaptations from inorganic to organic materials underscore the need for steering logics that address not just statistical shifts but also physicochemical contextual divergences [6, 7, 17].

This discussion extends to the role of architectural innovations, where graph-based models facilitate reusability but require boundary interpretations to avoid overgeneralization [25, 26, 29]. The RBA's pipeline dynamics suggest that feedback mechanisms can mitigate these risks, aligning with studies on deep transfer frameworks that emphasize iterative refinement [4, 10, 11].

However, challenges persist in multimodal integration, where epistemic risks from disparate data sources complicate cross-class applications, calling for enhanced uncertainty handling [18-20, 22].

In the context of discovery workflows, the framework's insights highlight trade-offs in closed-loop systems, where reusability boundaries influence the coupling of high-throughput simulations with experimental validations [1, 23, 24, 32]. This implies a shift toward adaptive infrastructures that dynamically bound transfers, fostering resilience in inverse design tasks [9, 16, 28]. Literature on foundation models further supports this, indicating that broad pre-training benefits from boundary-aware fine-tuning to maintain applicability across diverse classes [13, 30, 31].

Moreover, the RBA prompts consideration of broader field implications, such as in sustainable materials development, where cross-class reusability accelerates innovation by leveraging shared knowledge ecosystems [2, 8, 15]. Yet, potential pitfalls include entrenched biases from source domains, necessitating interpretive safeguards in computational steering [14, 21, 27]. By synthesizing these elements, the discussion affirms the RBA's value in bridging conceptual gaps, promoting integrative approaches that enhance the robustness of materials AI without empirical dependencies.

Conclusion

The Reusability Boundary Architecture (RBA) provides a novel conceptual scaffold for understanding the limits and potentials of transfer learning across materials classes in computational and data-driven engineering. Through its layered structure, pipelines, and steering logics, the framework interprets the systemic interactions that define reusability, offering insights into representation alignments, epistemic trade-offs, and discovery dynamics. This contributes to refining workflows in materials informatics, from multimodal data handling to autonomous systems, ultimately supporting accelerated materials innovation. Future explorations could extend these boundaries to emerging paradigms, reinforcing the field's computational foundations.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 10 Mar 2024 Revised: 15 Apr 2024 Accepted: 16 May 2024
Published online: 18 September 2024

References

Choudhary K, DeCost B, Chen Y, Hattrick-Simpers J, Chen Z, Chuang Y, et al. Recent advances and applications of deep learning methods in materials science. npj Comput Mater. 2022;8(1):59.

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.

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.

Chen C, Ong SP. AtomSets as a hierarchical transfer learning framework for small and large materials datasets. npj Comput Mater. 2021;7(1):173.

Jha D, Choudhary K, Tavazza F, Liao W-k, Choudhary A, Campbell C, et al. Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning. Nat Commun. 2019;10(1):5316.

Li X, Zhang Y, Zhao H, Burkhart C, Brinson LC, Chen W. A Transfer learning approach for microstructure reconstruction and structure-property predictions. Sci Rep. 2018;8(1):13461.

Li Y, Zhu R, Wang Y, Feng L, Liu Y. Center-environment deep transfer machine learning across crystal structures: From spinel oxides to perovskite oxides. npj Comput Mater. 2023;9(1):109.

Ren C, Li H, Dai Q, Wang R. Small data machine learning in materials science. npj Comput Mater. 2023;9(1):42.

Jiang L, Zhang Z, Hu H, He X, Fu H, Xie J. A rapid and effective method for alloy materials design via sample data transfer machine learning. npj Comput Mater. 2023;9(1):26.

Chang R, Wang YX, Ertekin E. Towards overcoming data scarcity in materials science: Unifying models and datasets with a mixture of experts framework. npj Comput Mater. 2022;8(1):242.

Feng S, Fu H, Zhou H, Wu Y, Lu Z, Dong H. A general and transferable deep learning framework for predicting phase formation in materials. npj Comput Mater. 2021;7(1):10.

Gupta V, Choudhary K, DeCost B, Tavazza F, Campbell C, Liao W-k, et al. Structure-aware graph neural network based deep transfer learning framework for enhanced predictive analytics on diverse materials datasets. npj Comput Mater. 2024;10(1):1.

Devi R, Butler KT, Sai Gautam G. Optimal pre-train/fine-tune strategies for accurate material property predictions. npj Comput Mater. 2024;10(1):300.

Goetz A, Durmaz AR, Müller M, Thomas A, Britz D, Kerfriden P, et al. Addressing materials' microstructure diversity using transfer learning. npj Compu Mater. 2022;8(1):27.

Yang F, Zhao W, Ru Y, Lin S, Huang J, Du B, et al. Transfer learning enables the rapid design of single crystal superalloys with superior creep resistances at ultrahigh temperature. npj Comput Mater. 2024;10(1):149.

Jiang C, He H, Guo H, Zhang X, Han Q, Weng Y, et al. Transfer learning guided discovery of efficient perovskite oxide for alkaline water oxidation. Nat Commun. 2024;15(1):6301.

Kaya M, Hajimirza S. Using a novel transfer learning method for designing thin film solar cells with enhanced quantum efficiencies. Sci Rep. 2019;9(1):5034.

Bets KV, O'Driscoll PC, Yakobson BI. Physics-inspired transfer learning for ML-prediction of CNT band gaps from limited data. npj Comput Mater. 2024;10(1):66.

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

Tan SIP, Liu Z, Ren Z, Chen H, Yu K, Yang S, et al. Tackling data scarcity with transfer learning: a case study of thickness characterization from optical spectra of perovskite thin films. Digit Discov. 2023;2(3):805-14.

Chen Z, Li D, Liu J, Gao K. Application of Gaussian processes and transfer learning to prediction and analysis of polymer properties. Comput Mater Sci. 2023;216:111859.

Schmidt J, Hoffmann N, Marques MAL. Transfer learning on large datasets for the accurate prediction of material properties. Digit Discov. 2023;2(5):1368-79.

Pettersson L, Verdozzi C. Crystal graph attention networks for the prediction of stable materials. Sci Adv. 2021;7(49):eabi7948.

Bonatti C, Mohr D. One for all: Universal material model based on minimal state-space neural networks. Sci Adv. 2021;7(26):eabf3658.

Chen X, Lu S, Chen Q, Zhou Q, Wang J. From bulk effective mass to 2D carrier mobility accurate prediction via adversarial transfer learning. Nat Commun. 2024;15(1):5391.

Lee J, Asahi R. Transfer learning for materials informatics using crystal graph convolutional neural network. Comput Mater Sci. 2021;190:110314.

Na GS. One-shot heterogeneous transfer learning from calculated crystal structures to experimentally observed materials. Comput Mater Sci. 2024;235:112791.

Ferrari BS, Manica M, Giro R, Steiner MB. Predicting polymerization reactions via transfer learning using chemical language models. npj Comput Mater. 2024;10(1):119.

Smith JS, Zubatyuk R, Nebgen B, Lubbers N, Barros K, Roitberg AE, et al. Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network. Sci Adv. 2019;5(8):eaav6490.

Pathrudkar S, Thiagarajan P, Agarwal S, Banerjee AS, Ghosh S. Electronic structure prediction of multi-million atom systems through uncertainty quantification enabled transfer learning. npj Comput Mater. 2024;10(1):175.

Zhang X, Ye G, Wen C, Bi Z. Transfer learning for predicting reorganization energy. Comput Mater Sci. 2023;228:112361.

Gongora AE, Snapp KL, Perry W, Li TJ, Tran TM, Pogue R, et al. Designing lattices for impact protection using transfer learning. Matter. 2022;5(9):2829-46.