

ORIGINAL RESEARCH

Open access

Perspective: Bridging ML and CALPHAD for Multicomponent Alloy Design — A Position on Hybrid Workflows

Emily Johnson^{1*}, Robert Smith¹, Laura Brown², Kevin Miller¹

Abstract

The design of multicomponent alloys, including high-entropy and complex concentrated systems, remains constrained by the methodological limits of existing computational frameworks. CALPHAD provides a physically rigorous basis for thermodynamic prediction, yet its scalability deteriorates sharply with increasing compositional complexity. Machine learning offers complementary strengths in navigating high-dimensional spaces and extracting latent structure from sparse data, but its predictions frequently lack thermodynamic consistency. This work advances the position that neither paradigm is sufficient in isolation and that their integration constitutes a necessary evolution in computational alloy design. A unified hybrid framework is proposed in which machine learning augments CALPHAD through accelerated parameterization, while thermodynamic constraints are embedded directly within learning architectures to enforce physical validity. This integration extends to iterative, closed-loop workflows in which active learning coordinates prediction, equilibrium evaluation, and targeted data acquisition. Such coupling reconciles the extrapolative capacity of data-driven models with the interpretability and stability of thermodynamic formalisms, thereby addressing the characteristic failure modes of each approach. The argument is situated within the rapid expansion of multicomponent alloy research, where the combinatorial growth of composition spaces has exposed the limitations of standalone methods. By aligning statistical inference with thermodynamic structure, hybrid workflows enable scalable and physically grounded exploration of complex systems. Realizing this potential requires interoperable computational infrastructure, open thermodynamic data, and validation strategies that prioritize both predictive accuracy and consistency with phase equilibria. Under these conditions, hybrid ML–CALPHAD approaches establish a coherent pathway toward predictive, physics-informed alloy design.

Keywords Machine learning, High-entropy alloys, CALPHAD, Hybrid workflows, Multicomponent alloys, Thermodynamic modeling

*Correspondence:

Emily Johnson
emily.johnson@gmail.com

¹ Department of Computational Materials Data Science, Faculty of Engineering, University of Toronto, Toronto, Canada

² Department of Data-Driven Materials Innovation, Faculty of Engineering, McGill University, Montreal, Canada

Introduction

The materials science community stands at a critical juncture in the design of multicomponent alloys. For more than half a century, the CALPHAD method has served as the cornerstone of thermodynamic modeling, enabling engineers to predict phase equilibria, phase fractions, and transformation temperatures across a wide range of engineering alloys [1, 2]. Yet the emergence of high-entropy alloys and other complex concentrated alloys—systems typically containing five or more principal elements—has exposed fundamental limitations in the traditional CALPHAD paradigm [3, 4]. At the same time, the explosive growth of machine learning in materials science has offered tantalizing new capabilities for navigating enormous composition spaces [5, 6].

This position paper asserts a clear stance: machine learning and CALPHAD are not competitors. They are complementary tools that must be bridged into hybrid workflows if meaningful progress is to be made in multicomponent alloy design. CALPHAD provides physically grounded interpolation within known thermodynamic space [1]. Machine learning provides pattern recognition and high-dimensional extrapolation [5, 6]. Neither alone can handle the complexity of systems with five or more principal elements [7, 8]. The community must therefore invest in hybrid workflows, shared infrastructure, and validation standards [9].

The position matters now because the composition space of multicomponent alloys is astronomically large. Even a modest five-element system at 10 at.% resolution yields more than 10^6 distinct compositions; six- and seven-element systems quickly reach tens of millions. Gorsse and Tancret [1] has long emphasized CALPHAD's strength in interpolation within calibrated ranges, yet extrapolation to new multicomponent regimes remains unreliable without extensive re-parameterization. Liu and co-workers [5] reviewed the rapid adoption of machine learning for high-entropy alloys and highlighted both its promise and its lack of thermodynamic grounding. Standalone machine learning models can identify promising regions but cannot guarantee that predicted energies correspond to physically realizable phase diagrams [7, 10].

Recent work by Zeng and colleagues [11] demonstrated that combining CALPHAD assessments with machine learning can reveal high-fidelity phase selection rules that neither approach uncovers independently. Similarly, Zou *et al.* [12] showed that integrating machine learning with CALPHAD accelerates the discovery of low-modulus near- β -Ti alloys. These studies, among others published between 2017 and 2024 [3, 13–16], illustrate that hybrid thinking is already producing results, yet the field still lacks a coherent framework, shared infrastructure, and community standards.

The proposed hybrid workflows operate at three integration levels of increasing sophistication. At the simplest level, machine learning accelerates the generation of CALPHAD parameters [17, 18]. At the intermediate level, CALPHAD thermodynamic constraints are imposed directly on machine learning models. At the most advanced level, iterative active-learning loops couple predictions, equilibrium calculations, and targeted first-principles data generation [19, 20]. Each level addresses specific gaps: the parameterization bottleneck of CALPHAD, the physical inconsistency of machine learning, and the data sparsity that limits both.

This position is not merely technical; it is strategic. Multicomponent alloys promise unprecedented combinations of strength, ductility, corrosion resistance, and high-temperature performance. Realizing that promise at industrial scales requires predictive tools that scale with complexity. Hybrid ML-CALPHAD workflows represent the only viable path forward. The remainder of this paper articulates the strengths and weaknesses of each paradigm, justifies the necessity of hybridization, details the proposed workflows, and offers concrete recommendations for the community.

CALPHAD

The CALPHAD method remains the most mature and widely adopted framework for thermodynamic modeling of alloys. Its core strength lies in its physical foundation: phase equilibria are computed from Gibbs free energy descriptions that are parameterized against experimental and first-principles data. Gorsse and Tancret [1] has eloquently described how CALPHAD has guided the development of advanced materials for decades by providing reliable interpolation within composition and temperature ranges where databases have been calibrated.

CALPHAD excels at handling multicomponent phase diagrams once the necessary binary, ternary, and higher-order interaction parameters are established. It naturally incorporates thermodynamic consistency—convexity of free energy surfaces, common-tangent constructions for phase equilibria, and temperature-dependent heat capacities. Extensive commercial and open databases have accumulated over forty years, covering thousands of binary and ternary systems. High-throughput CALPHAD implementations, as reviewed by Ghassemali [4], further accelerate screening of large

composition spaces [9]. Li and co-workers [3] successfully employed CALPHAD-aided design to optimize thermal stability in a refractory high-entropy alloy, demonstrating the method's continuing utility even in complex systems [21].

Yet CALPHAD's weaknesses become prohibitive precisely when multicomponent alloys are the target. Extrapolation beyond calibrated ranges is unreliable because interaction parameters derived from lower-order systems do not necessarily extend to quinary or higher compositions. Parameterization itself is labor-intensive; each new subsystem requires careful assessment against experimental or DFT data. For high-entropy alloys containing five or more principal elements, the number of required subsystems grows combinatorially. Databases for many technologically relevant multicomponent spaces simply do not exist, and building them from scratch can take years.

Deffrennes and colleagues [13] noted that even state-of-the-art CALPHAD assessments struggle with phase-diagram prediction in unexplored regions. Cao *et al.* [22] attempted to predict enthalpies of mixing using machine learning augmentation of CALPHAD databases, underscoring the growing recognition that traditional parameterization cannot keep pace. Bansal and others [16] have shown that high-throughput CALPHAD alone is insufficient for accelerated design of high-entropy alloys precisely because of the parameterization bottleneck.

In short, CALPHAD provides trustworthy interpolation and thermodynamic rigor within known spaces but collapses under the combinatorial explosion of multicomponent systems. Its strengths are foundational; its weaknesses are structural. Hybridization is therefore not an enhancement—it is a necessity.

Machine Learning

Machine learning has emerged as a powerful complement to physics-based modeling by learning complex patterns directly from data. Its primary strength in the alloy-design context is the ability to handle high-dimensional composition spaces without requiring explicit thermodynamic models. Liu *et al.* [5] surveyed the rapid progress of machine learning for high-entropy alloys and highlighted its capacity to screen millions of compositions in seconds once trained.

Recent studies illustrate this capability. Zeng and colleagues [11] used machine learning to uncover phase-selection rules that traditional CALPHAD assessments had missed. Zou *et al.* [12] integrated machine learning with CALPHAD to accelerate discovery of low-modulus β -Ti alloys. Wang *et al.* [23] reviewed recent progress integrating DFT, CALPHAD, and machine learning and further demonstrates how data-driven models can suggest promising regions that would be impractical to explore manually. Batzner and co-workers [24] showed that equivariant graph neural networks can generate accurate interatomic potentials with remarkably small training sets, opening avenues for machine-learning-augmented thermodynamic modeling.

Machine learning is also fast. Once trained, predictions are near-instantaneous, enabling rapid iteration in design campaigns. It naturally incorporates diverse data sources—experimental measurements, DFT energies, and even literature-extracted values—without requiring manual parameterization of interaction terms [14, 15].

Nevertheless, machine learning's weaknesses are equally pronounced. Most models are not constrained by thermodynamics. Predicted energies may violate Gibbs free energy convexity, leading to impossible phase equilibria. Phase diagrams generated from purely data-driven formation energies are frequently inconsistent with the common-tangent rule. The review on “Machine learning for CALPHAD” [7] and the hybrid approach paper [8] both emphasize that unconstrained machine learning can produce thermodynamically invalid predictions.

Data requirements constitute another critical weakness. Although machine learning can extrapolate, its reliability outside the training distribution is uncertain. Gao *et al.* [25] and Alpaydin [19] both acknowledge that sparse DFT datasets for high-entropy alloys limit generalization. Without thermodynamic grounding, models risk overfitting to artifacts rather than learning true underlying physics [10].

Thus machine learning offers speed and scalability but lacks the physical consistency that CALPHAD provides. Its strengths are exploratory; its weaknesses are foundational. Only hybridization can retain the best of both worlds.

Why Hybrid Is Necessary for Multicomponent Alloys

Five compelling reasons demonstrate that hybrid ML-CALPHAD workflows are not optional but essential for multicomponent alloy design.

First, the composition space is simply too large for CALPHAD alone. A quinary alloy at 5 at.% resolution already exceeds 10^7 compositions. Traditional parameterization cannot scale [4, 9, 16]. Second, thermodynamic constraints are too important to ignore in machine learning. Unconstrained models can predict negative curvature in free-energy surfaces, implying spontaneous phase separation that violates stability criteria [7, 8]. Third, data remain too sparse. Even large DFT campaigns generate only hundreds of structures for any given high-entropy system, insufficient for purely data-driven models to learn reliable thermodynamics [19, 25].

Fourth, extrapolation is unavoidable yet risky for both approaches. CALPHAD extrapolation fails outside assessed subsystems; machine learning extrapolation lacks uncertainty quantification tied to physics [10]. Hybridization allows cross-validation: machine learning proposes candidates, CALPHAD enforces consistency [11, 12, 14]. Fifth, different length and time scales demand different tools. Atomistic machine-learning potentials feed into CALPHAD for microstructure-scale predictions. Without integration, information is lost between scales [23, 24].

Empirical evidence from the literature supports each point. Li et al. [3] showed that CALPHAD alone could guide a TiZrHfNb alloy but required machine-learning augmentation for optimal thermal stability [21]. Deffrennes [13] demonstrated that machine-learning classification improves phase-diagram prediction but still needs CALPHAD equilibrium calculations for validation. Ghassemali [4], Ruan et al. [26], and Alpaydin [19] all conclude that neither method suffices independently.

The hybrid imperative is therefore clear: multicomponent alloys expose the limits of both paradigms simultaneously. Only integrated workflows can deliver physically consistent, computationally scalable, and experimentally validated predictions.

Proposed Hybrid Workflows

Three integration levels offer a practical roadmap from incremental improvement to transformative capability.

The three hierarchical integration levels, their functional roles, and their progressively increasing design capabilities are structurally formalized in Table 1.

Table 1. Hierarchical Integration Levels in Hybrid ML–CALPHAD Workflows: Functional Roles, Information Flow, and Design Capabilities

Integration Level	Core Functional Role	Information Flow Structure	Constraint Enforcement	Data Dependency	Design Capability
Level 1: ML-Accelerated CALPHAD	Accelerate thermodynamic parameter generation	ML → CALPHAD (unidirectional)	External (CALPHAD only)	Moderate (requires subsystem data)	Rapid database expansion
Level 2: CALPHAD-Constrained ML	Enforce thermodynamic	ML ↔ Constraint Layer (embedded)	Internal (loss function penalties)	Moderate to high	Physically valid property prediction

	consistency in ML predictions				
Level 3: Iterative Hybrid	Close prediction–validation–data acquisition gap	Sequential refinement across components	Dynamic (evaluation-driven)	Adaptive (active learning)	Autonomous alloy discovery
Cross-Level Property	Increasing integration depth	Increasing bidirectional coupling	Increasing physical fidelity	Increasing efficiency	Increasing scalability

Level 1—ML-accelerated CALPHAD—uses machine learning to predict binary and ternary interaction parameters that are then fed directly into standard CALPHAD solvers. This approach dramatically reduces parameterization time for new multicomponent systems. Recent work by Zhang and colleagues [17] on Cu–Ni–Sn–Al alloys and by Feng Liu [18] on NiCoCrAl eutectic high-entropy alloys illustrates early successes of this level.

Level 2—CALPHAD-constrained machine learning—embeds thermodynamic consistency directly into model training. Loss functions are augmented with penalties for non-convex free energies or violations of phase-equilibrium conditions. The resulting models produce formation energies that are guaranteed to yield physically realizable phase diagrams when passed to a CALPHAD engine [7, 8].

Level 3—tightly coupled iterative hybrid—represents the most powerful architecture. The workflow begins with sparse DFT data for key compositions. A machine-learning model is trained to predict energies across the full composition space. CALPHAD then computes the phase diagram from these energies. Discrepancies or high-uncertainty regions are automatically flagged. Targeted DFT calculations are performed in those regions, the new data are added to the training set, and the cycle repeats. Active learning ensures data efficiency. Alpaydin [19] and Odetola et al. [20] both point toward this iterative paradigm.

A conceptual diagram (Figure 1) illustrates the three levels. Figure 1 illustrates the three-level hybrid ML–CALPHAD architecture proposed for scalable, thermodynamically consistent multicomponent alloy design.

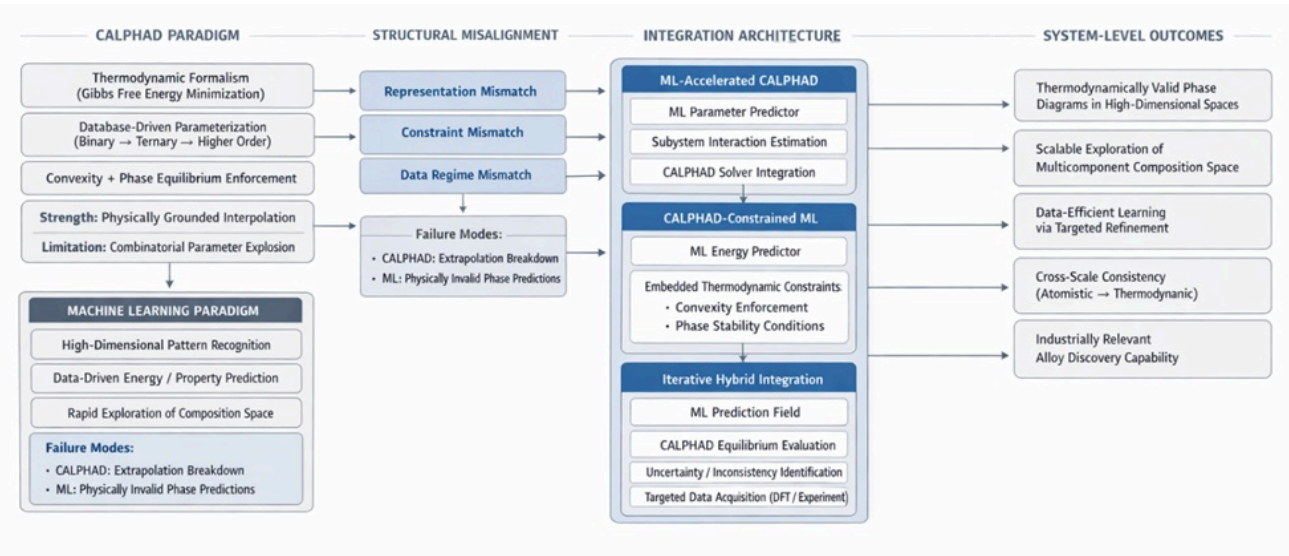


Figure 1. Hierarchical Architecture of Complementarity, Constraint Imposition, and Iterative Integration in Hybrid ML–CALPHAD Workflows for Multicomponent Alloy Design

Infrastructure requirements are non-negotiable. Common data formats must enable seamless exchange between machine-learning frameworks and CALPHAD software such as Thermo-Calc and OpenCalphad. Standardized APIs are needed. Community benchmark datasets—curated mixtures of experimental, DFT, and CALPHAD-derived values—will allow objective comparison of hybrid implementations.

These workflows are not speculative. Elements of each level already appear in the literature from 2017 to 2024 [11, 12, 17, 18]. The community's task is to formalize, standardize, and scale them.

Objections and Responses

Addressing persistent objections to hybrid ML–CALPHAD integration

Despite accumulating support for hybrid ML–CALPHAD workflows, persistent skepticism within the materials science community reflects deeper epistemic and institutional frictions. The perceived disconnect between CALPHAD and ML communities signals not an intrinsic incompatibility but a coordination failure that can be mitigated through shared infrastructures and joint problem framings. Emerging high-throughput initiatives already demonstrate how collaborative mechanisms—particularly interoperable data repositories and co-developed benchmarks—reconfigure this divide into a site of productive exchange [4, 9]. Embedding such interaction within standardized APIs and competitive benchmarking environments further institutionalizes cross-community engagement, transforming disciplinary separation into complementary specialization [19, 20].

Concerns regarding the limited physical grounding of machine learning reveal a more substantive tension around model legitimacy in scientific discovery. This critique, however, underscores the necessity of integrating thermodynamic structure directly into learning architectures. Rather than opposing paradigms, ML and CALPHAD occupy distinct yet mutually reinforcing epistemic roles: the former excels in extracting latent patterns, while the latter enforces consistency with established thermodynamic laws. CALPHAD-constrained learning frameworks operationalize this synthesis by embedding equilibrium and convexity constraints into training objectives, thereby aligning predictive flexibility with physical plausibility. Evidence from physics-informed architectures indicates that such constraints can preserve predictive accuracy while enforcing fundamental invariances, suggesting that extending these principles to thermodynamic consistency is both methodologically coherent and practically attainable [7, 8, 24].

The proprietary nature of many CALPHAD databases introduces an additional structural barrier, yet this limitation is increasingly contingent rather than absolute. The expansion of open-source thermodynamic repositories signals a gradual shift toward more accessible data ecosystems, particularly in high-throughput contexts where reproducibility and scalability are paramount. OpenCalphad and NIST initiatives exemplify this trajectory, with accessible datasets enabling broader participation and accelerating iterative model refinement [4, 9]. Under these conditions, hybrid workflows act as a catalyst for further openness by incentivizing shared, version-controlled parameter infrastructures that support both thermodynamic assessment and machine learning integration [20].

Apprehensions about workflow complexity often conflate conceptual integration with practical implementation. In practice, the modular architecture of hybrid approaches allows incremental adoption, where early-stage integration requires minimal deviation from existing computational pipelines. The feasibility of ML-accelerated parameter estimation within established CALPHAD solvers illustrates how complexity emerges only in response to expanded analytical demands rather than as an inherent prerequisite. Empirical studies in multicomponent alloy systems demonstrate that even limited integration yields measurable gains in predictive efficiency and design insight, reinforcing the argument that complexity is situational rather than prohibitive [17, 18, 27, 28].

Skepticism toward the accuracy of DFT data in CALPHAD contexts reflects a legitimate concern regarding error propagation across scales. Hybrid workflows, however, reframe this limitation as an opportunity for adaptive data integration. By incorporating experimental observations alongside first-principles calculations and prioritizing uncertain

regions through active-learning strategies, these systems dynamically allocate computational and experimental resources. Iterative refinement processes thereby reduce dependence on any single data source while enhancing overall model robustness, as demonstrated in recent uncertainty-driven thermodynamic studies [19, 20].

Viewed through this lens, each objection delineates not a boundary but a design constraint that hybrid frameworks are uniquely positioned to address. The integration of ML and CALPHAD does not demand disciplinary convergence in the traditional sense; it instead relies on orchestrating their distinct strengths within a coherent computational architecture. Progress therefore hinges as much on organizational alignment as on technical innovation, with the trajectory toward hybridization already discernible across recent developments.

Relation to adjacent theoretical and methodological positions

This position extends and consolidates several influential strands of prior work while sharpening the emphasis on thermodynamic integration in multicomponent alloy design. It resonates with the physics-constrained machine learning paradigm advanced by Batzner *et al.* [24], where embedding symmetry considerations within neural architectures enhances both efficiency and predictive fidelity. Translating this principle into the thermodynamic domain shifts the focus from geometric invariance to energy landscape consistency, where enforcing Gibbs free energy convexity and phase-equilibrium conditions serves an analogous role in constraining model behavior [7, 8]. In both cases, the rejection of purely black-box formulations reflects a broader epistemological commitment to embedding domain knowledge within data-driven systems.

A parallel alignment emerges with multi-fidelity modeling strategies, where the interplay between computational efficiency and physical accuracy is explicitly managed. Within this framework, CALPHAD functions as a rapid, lower-fidelity evaluator that enables large-scale exploration, while DFT and targeted experiments provide higher-fidelity anchors that refine model predictions. Iterative hybrid workflows operationalize this hierarchy by selectively invoking expensive calculations in regions of high uncertainty, thereby optimizing resource allocation and accelerating convergence. Empirical demonstrations highlight how thermodynamic screening can guide high-fidelity data acquisition, reinforcing the strategic coupling of modeling scales [4, 13, 14].

The integration of active learning further deepens this synthesis by introducing adaptive feedback mechanisms into the thermodynamic loop. Uncertainty-driven sampling, as demonstrated in phase-diagram prediction studies, becomes substantially more effective when coupled with thermodynamic validation criteria [19]. Situating active learning within a closed-loop architecture—where machine learning proposes candidate energies, CALPHAD evaluates equilibrium states, and discrepancies trigger targeted first-principles calculations—addresses a key limitation of earlier approaches that relied on uncertainty metrics alone [23]. This configuration transforms model refinement into a physically grounded, iterative process.

By unifying physics-constrained learning, multi-fidelity modeling, and active learning under the principle of thermodynamic consistency, the hybrid ML–CALPHAD framework consolidates previously fragmented advances into a coherent methodological paradigm. The resulting architecture not only clarifies the conceptual relationships among these approaches but also provides a concrete pathway for their joint application in multicomponent alloy design.

Recommendations for institutional and community alignment

Realizing this paradigm requires coordinated shifts across the research ecosystem, particularly in how knowledge is produced, shared, and validated. For CALPHAD practitioners, greater openness in thermodynamic parameterization and software interoperability becomes essential to enabling seamless integration with machine learning environments. Machine-readable data formats and accessible solver interfaces lower the barrier to entry while facilitating reproducibility and cross-platform experimentation. Engagement in benchmark-driven evaluation further situates thermodynamic modeling within a comparative, community-wide context that accelerates methodological convergence [4, 9].

From the perspective of machine learning research, meaningful participation in this integration demands a deeper engagement with thermodynamic principles. Embedding constraints such as convexity and equilibrium conditions directly into model training shifts validation from isolated property prediction toward system-level consistency. Treating CALPHAD not as a post hoc verification step but as an intrinsic component of the modeling pipeline redefines the role of ML within materials discovery, aligning predictive performance with physical interpretability [8, 10].

Institutional support plays a decisive role in sustaining this transition. Funding structures that prioritize interdisciplinary collaboration and the development of shared thermodynamic–ML infrastructures can accelerate progress while reducing fragmentation. Investments in open databases and standardized evaluation protocols create the conditions under which hybrid approaches can scale beyond isolated case studies, particularly when aligned with initiatives emphasizing integrated, closed-loop methodologies [19, 20].

Editorial policies likewise shape the trajectory of the field by establishing norms around validation and data transparency. Emphasizing thermodynamic consistency in ML-driven alloy studies and requiring the public availability of both predictive models and associated CALPHAD inputs elevates the evidentiary standard for publication. Such measures not only enhance reproducibility but also incentivize the development of genuinely integrative approaches over parallel, non-interacting methodologies [11, 12].

Existing studies already foreshadow this transition, with combined CALPHAD–ML strategies demonstrating tangible gains in alloy design efficiency and predictive capability [3, 11, 12, 14, 15]. The challenge now lies in consolidating these advances into a shared, extensible ecosystem that supports cumulative progress rather than isolated innovation.

A call to action for hybrid ML–CALPHAD realization

The transition toward hybrid ML–CALPHAD workflows has moved beyond conceptual advocacy to practical necessity. Near-term progress depends on establishing institutional and technical foundations that enable sustained collaboration, including coordinated community efforts, shared benchmark datasets, and interoperable computational interfaces linking modern ML frameworks with thermodynamic solvers [4, 9, 19]. Such developments formalize the integration process and reduce reliance on ad hoc implementations.

Figure 2 summarizes the infrastructure, validation, and community-alignment requirements needed to convert hybrid ML–CALPHAD workflows from isolated demonstrations into standardized materials-design practice.

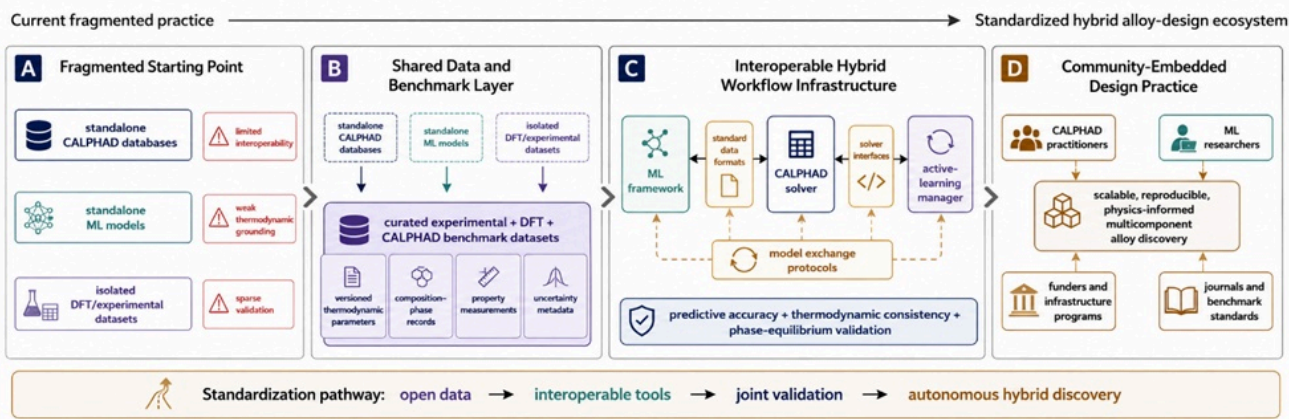


Figure 2. Implementation Roadmap for Standardizing Hybrid ML–CALPHAD Workflows in Multicomponent Alloy Discovery

As the field matures, emphasis will shift toward scalability and standardization, particularly through the development of community-curated thermodynamic parameter databases and rigorous validation protocols that jointly assess predictive accuracy and thermodynamic consistency. Embedding hybrid workflows within broader materials design platforms further extends their accessibility, allowing researchers to leverage advanced methodologies without requiring deep specialization in either CALPHAD or machine learning [8, 20].

Over longer horizons, the convergence of these elements points toward autonomous, physics-informed alloy design systems in which machine learning, CALPHAD, and first-principles calculations operate within a fully integrated feedback loop. Such systems enable continuous refinement across scales, from atomic interactions to macroscopic phase behavior, while lowering barriers to entry for exploring complex compositional spaces. The cumulative effect is a democratization of alloy design capabilities alongside a substantial increase in exploratory efficiency.

Prior work has already outlined the contours of this trajectory, from early conceptual mappings of hybrid opportunities to iterative demonstrations of feasibility [5, 8, 19, 20]. The remaining challenge is organizational rather than conceptual: translating isolated successes into standardized practice. The direction is established; the pace of realization now depends on the collective willingness to institutionalize hybrid methodologies within the core of materials discovery.

Conclusion

Machine learning and CALPHAD are complementary, not competitors. Neither alone is sufficient for multicomponent alloy design. CALPHAD delivers physically grounded interpolation within known thermodynamic space; machine learning supplies pattern recognition and high-dimensional extrapolation. Their strengths and weaknesses are mirror images: where one excels the other falters, and vice versa.

This position paper has articulated three practical integration levels—ML-accelerated CALPHAD, CALPHAD-constrained machine learning, and tightly coupled iterative hybrids—that together close the critical gaps exposed by high-entropy and complex concentrated alloys. It has addressed common objections, situated the hybrid framework within related positions on physics-constrained ML, multi-fidelity modeling, and active learning, and offered concrete recommendations for practitioners, researchers, funders, and editors.

The call to action is urgent yet achievable: establish working groups, release benchmark datasets, develop open APIs, and institutionalize thermodynamic consistency checks. The literature from 2017 to 2024 already contains compelling proof-of-concept studies. The infrastructure and community standards proposed here will transform those isolated successes into a robust, scalable methodology.

The future of alloy design is hybrid. By deliberately bridging ML and CALPHAD, the materials community can move beyond the combinatorial impasse of multicomponent systems and unlock the extraordinary property combinations that high-entropy and complex alloys promise. The time to invest in shared infrastructure, validation protocols, and interdisciplinary collaboration is now. Only through deliberate integration will computational materials engineering fulfill its promise for the alloys of tomorrow.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 21 Nov 2023

Revised: 27 Jan 2024

Accepted: 05 May 2024

Published online: 18 July 2024

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

Gorsse S, Tancrét F. Current and emerging practices of CALPHAD toward the development of high entropy alloys and complex concentrated alloys. *J Mater Res.* 2018;33(19):2899–923.

<https://doi.org/10.1557/jmr.2018.152>.

Ryzhov VN, Tareyeva EE, Fomin YD, Tsiok EN. Complex phase diagrams of systems with isotropic potentials: Results of computer simulations. *Phys Usp.* 2020;63(5):417–39.

<https://doi.org/10.3367/UFNe.2018.04.038417>.

Li T, Wang S, Fan W, Lu Y, Wang T, Li T, et al. CALPHAD-aided design for superior thermal stability and mechanical behavior in a TiZrHfNb refractory high-entropy alloy. *Acta Mater.* 2023;246:118728.

<https://doi.org/10.1016/j.actamat.2023.118728>.

Ghassemali E, Conway PLJ. High-throughput CALPHAD: A powerful tool towards accelerated metallurgy. *Front Mater.* 2022;9:889771.

<https://doi.org/10.3389/fmats.2022.889771>.

Liu X, Zhang J, Pei Z. Machine learning for high-entropy alloys: Progress, challenges and opportunities. *Prog Mater Sci.* 2023;131:101018.

<https://doi.org/10.1016/j.pmatsci.2022.101018>.

Zhou ZH. Machine learning. Singapore: Springer; 2021.

<https://doi.org/10.1007/978-981-15-1967-3>.

Zhang C, Yang Y. The CALPHAD approach for HEAs: Challenges and opportunities. *MRS Bull.* 2022;47(2):158–67.

<https://doi.org/10.1557/s43577-022-00284-8>.

Taheri-Mousavi SM, Xu M, Hengsbach F, Houser C, Ge Z, Glaser B, et al. Additively manufacturable high-strength aluminum alloys with thermally stable microstructures enabled by hybrid machine learning-based design. *arXiv [Preprint]*. 2024.

<https://doi.org/10.48550/arXiv.2406.17457>.

van de Walle A, Sun R, Hong QJ, Kadkhodaei S. Software tools for high-throughput CALPHAD from first-principles data. *Calphad*. 2017;58:70-81.
<https://doi.org/10.1016/j.calphad.2017.05.005>.

Hsu HHH, Shen Y, Tomani C, Cremers D. What makes graph neural networks miscalibrated? *Adv Neural Inf Process Syst*. 2022;35:13775-86.

Zeng Y, Man M, Bai K, Zhang YW. Revealing high-fidelity phase selection rules for high entropy alloys: A combined CALPHAD and machine learning study. *Mater Des*. 2021;202:109532.
<https://doi.org/10.1016/j.matdes.2021.109532>.

Zou H, Tian YY, Zhang LG, Xue RH, Deng ZX, Lu MM, et al. Integrating machine learning and CALPHAD method for exploring low-modulus near- β -Ti alloys. *Rare Met*. 2024;43(1):309-23.
<https://doi.org/10.1007/s12598-023-02333-w>.

Deffrennes G, Terayama K, Abe T, Tamura R. A machine learning-based classification approach for phase diagram prediction. *Mater Des*. 2022;215:110497.
<https://doi.org/10.1016/j.matdes.2022.110497>.

Li W, Raman L, Debnath A, Ahn M, Lin S, Krajewski AM, et al. Design and validation of refractory alloys using machine learning, CALPHAD, and experiments. *Int J Refract Met Hard Mater*. 2024;121:106673.
<https://doi.org/10.1016/j.ijrmhm.2024.106673>.

Jha R, Chakraborti N, Diercks DR, Stebner AP, Ciobanu CV. Combined machine learning and CALPHAD approach for discovering processing-structure relationships in soft magnetic alloys. *Comput Mater Sci*. 2018;150:202-11.
<https://doi.org/10.1016/j.commatsci.2018.04.008>.

Bansal A, Kumar P, Yadav S, Hariharan VS, Rahul MR, Phanikumar G. Accelerated design of high entropy alloys by integrating high throughput calculation and machine learning. *J Alloys Compd*. 2023;960:170543.
<https://doi.org/10.1016/j.jallcom.2023.170543>.

Zhang W, Tang Y, Gao J, Zhang L, Ding J, Xia X. Determination of hardness and Young's modulus in fcc Cu-Ni-Sn-Al alloys via high-throughput experiments, CALPHAD approach and machine learning. *J Mater Res Technol*. 2024;30:5381-93.
<https://doi.org/10.1016/j.jmrt.2024.04.221>.

Liu F, Xiao X, Huang L, Tan L, Liu Y. Design of NiCoCrAl eutectic high entropy alloys by combining machine learning with CALPHAD method. *Mater Today Commun*. 2022;30:103172.
<https://doi.org/10.1016/j.mtcomm.2022.103172>.

Alpaydin E. *Machine learning*. Revised and updated ed. Cambridge (MA): MIT Press; 2021. 280 p.

Odetola PI, Babalola BJ, Afolabi AE, Anamu US, Olorundaisi E, Umba MC, et al. Exploring high entropy alloys: A review on thermodynamic design and computational modeling strategies for advanced materials applications. *Heliyon*. 2024;10(22):e39660.
<https://doi.org/10.1016/j.heliyon.2024.e39660>.

Guo L, Gu J, Gong X, Ni S, Song M. CALPHAD-aided design of high entropy alloy to achieve high strength via precipitate strengthening. *Sci China Mater*. 2020;63(2):288-99.
<https://doi.org/10.1007/s40843-019-1170-7>.

Cao X, Luo W, Liu H. A prediction model for CO₂/CO adsorption performance on binary alloys based on machine learning. *RSC Adv*. 2024;14(17):12235-46.
<https://doi.org/10.1039/D4RA00710G>.

Wang Y, Li K, Soisson F, Becquart CS. Combining DFT and CALPHAD for the development of on-lattice interaction models: The case of Fe-Ni system. *Phys Rev Mater*. 2020;4(11):113801.

<https://doi.org/10.1103/PhysRevMaterials.4.113801>.

Batzner S, Musaelian A, Sun L, Geiger M, Mailoa JP, Kornbluth M, et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat Commun.* 2022;13(1):2453.

<https://doi.org/10.1038/s41467-022-29939-5>.

Gao J, Zhong J, Liu G, Yang S, Song B, Zhang L, et al. A machine learning accelerated distributed task management system (Malac-Distmas) and its application in high-throughput CALPHAD computations aiming at efficient alloy design. *Adv Powder Mater.* 2022;1(1):100005.

<https://doi.org/10.1016/j.apmate.2021.09.005>.

Ruan J, Xu W, Yang T, Yu J, Yang S, Luan J, et al. Accelerated design of novel W-free high-strength Co-base superalloys with extremely wide γ/γ' region by machine learning and CALPHAD methods. *Acta Mater.* 2020;186:425-33.

<https://doi.org/10.1016/j.actamat.2020.01.004>.

Korotaev P, Yanilkin A. Steels classification by machine learning and CALPHAD methods. *Calphad.* 2023;82:102587.

<https://doi.org/10.1016/j.calphad.2023.102587>.

Paulus K. Combined CALPHAD and machine learning for property modelling [master's thesis]. Stockholm: KTH Royal Institute of Technology; 2020. Available from: <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-278149>