

ORIGINAL RESEARCH

Open access

Defining "Novel Material" in Generative Models: A Boundary between Chemical Validity and Synthesizability

Khaled Mahfouz^{1*}, Rania Abdelaziz¹, Sherif Adel², Dina Mostafa¹, Tamer Nabil², Reem Saad¹

Abstract

"Novel material" has become the central claim in nearly every paper introducing generative models for materials discovery. Yet the term is used with striking ambiguity. Authors routinely assert that their models have produced "novel materials" without clarifying whether this means the output lies outside the training distribution, satisfies basic chemical rules such as charge balance and realistic bond lengths, or meets the far stricter requirement of being synthesizable in a laboratory. This boundary/definitional article identifies three distinct meanings—distributional novelty, chemical validity, and synthesizability—and demonstrates how the current literature routinely conflates them. A generated crystal may be distributionally novel (absent from the training set) yet chemically invalid; it may be chemically valid yet lie far above the convex hull and therefore remain unsynthesizable. Such conflation leads to overclaims that inflate expectations and hinder reproducible progress in inverse design. This analysis maps the boundary conditions for each meaning and proposes a hierarchical operational definition with six explicit levels. The framework requires authors to report novelty percentages at every level rather than a single vague statistic. Distributional novelty marks the first filter, chemical validity the second, and synthesizability the third, with synthesizability itself subdivided into thermodynamic, dynamical, kinetic, and experimental realizability layers. The article further examines boundary cases, gray zones, and implications for model evaluation and benchmark design. By replacing ambiguous rhetoric with precise, multi-level reporting, the proposed definition establishes a shared language for generative models in materials engineering and prevents the overinterpretation of computational outputs as laboratory-ready discoveries. Adoption of this operational framework will sharpen claims, improve comparability across studies, and ultimately accelerate the translation of generative predictions into experimentally validated materials.

Keywords Generative models, Inverse design, Synthesizability, Chemical validity, Novel material, Distributional novelty

*Correspondence:

Khaled Mahfouz
khaled.mahfouz@gmail.com

¹ Department of Intelligent Materials Engineering, Faculty of Engineering, Mansoura University, Mansoura, Egypt

² Department of Computational Materials Analytics, Faculty of Engineering, Helwan University, Cairo, Egypt

Introduction

Every generative model paper for materials makes a claim: "We discovered novel materials" [1-3]. But what does "novel" actually mean? Does it mean the material is not in the training set? Does it mean the material is chemically valid—satisfying charge balance and reasonable bond lengths? Does it mean the material can be synthesized in a laboratory? The term is used ambiguously, conflating different concepts that carry very different scientific weight [4, 5]. A material can be absent from the training set yet chemically invalid. It can be chemically valid yet thermodynamically unstable and therefore unsynthesizable [6, 7]. This ambiguity is not merely semantic; it directly affects how the community evaluates progress in inverse design and how funding agencies and experimentalists interpret computational promises [8, 9].

Figure 1 illustrates that “novel material” is not a single property but a hierarchical filtering process in which distributional novelty, chemical validity, and multiple layers of synthesizability progressively constrain generative outputs, while misinterpretation at lower levels leads to systematic overclaim.

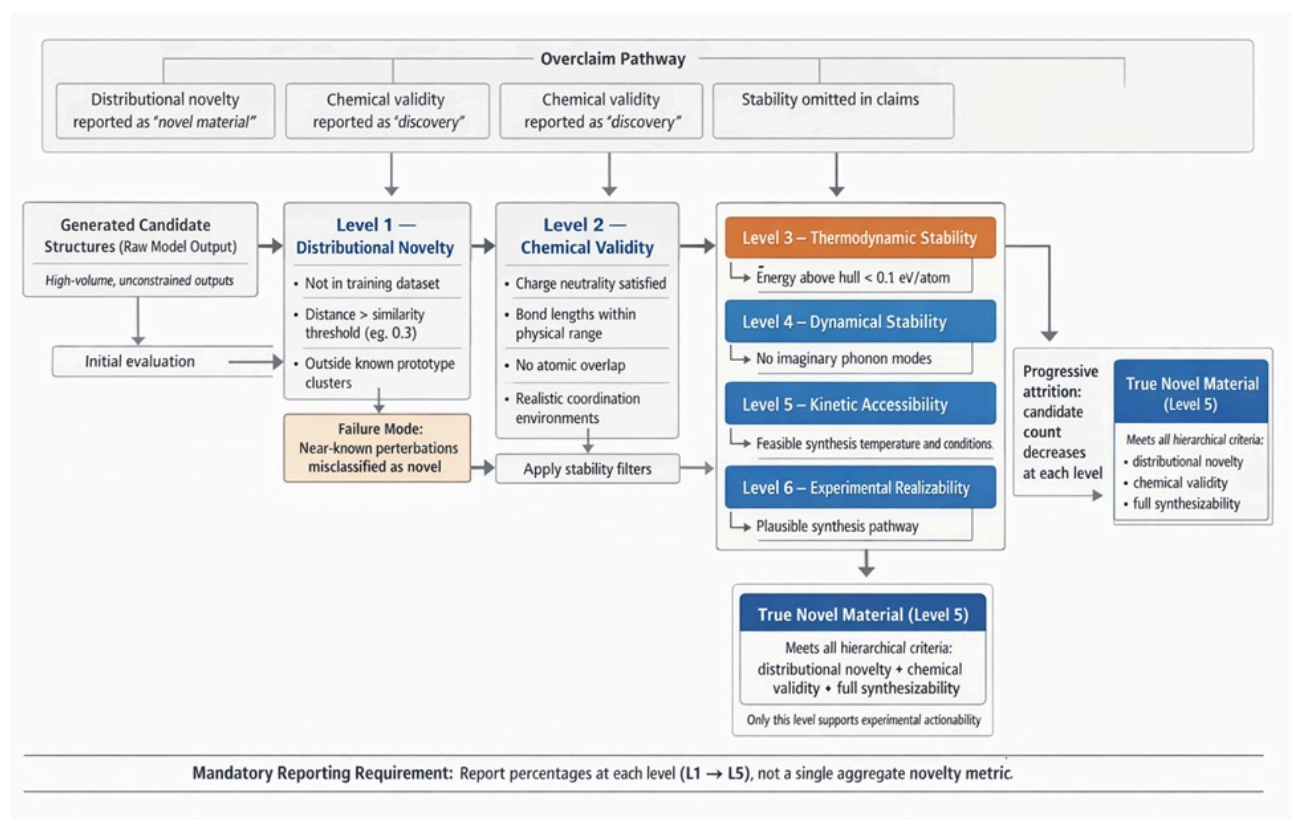


Figure 1. Hierarchical Failure of “Novel Material” Claims: From Distributional Novelty to Experimental Synthesizability

Recent reviews of generative approaches to inorganic crystal design [3, 4] highlight the explosion of diffusion-based, flow-based, and reinforcement-learning frameworks that claim to expand the known materials space [10–13]. Yet the same papers rarely define the threshold at which a generated structure crosses from “known” to “novel.” Some authors equate novelty with simple non-duplication in the training database, while others invoke chemical-validity checks without mentioning synthesizability [14, 15]. Still others declare novelty on the basis of an unreported structure without providing stability or pathway analysis [16]. This lack of consensus creates a situation in which two models reporting “95 % novel materials” may actually be describing entirely different phenomena. One model may have produced structures that are merely slight perturbations of training examples; another may have generated charge-balanced crystals that lie high above the convex hull; a third may have reached structures that satisfy thermodynamic and dynamical stability yet remain kinetically inaccessible.

The present boundary/definitional article addresses this gap by providing a rigorous conceptual analysis grounded exclusively in the peer-reviewed literature from 2017 to 2025. It distinguishes chemical validity from synthesizability and demonstrates that these are not interchangeable filters but sequential gates that must be applied hierarchically [6, 17]. The analysis draws on studies that have explicitly examined synthesizability prediction, chemical-validity metrics in crystal generation, and the evaluation of generative models for materials [7, 14, 15, 18]. By clarifying the boundary conditions for each meaning of novelty, the article supplies an operational definition that future authors can adopt verbatim.

The stakes are practical as well as philosophical. Experimental groups cannot pursue every computationally generated candidate. When generative models flood the literature with thousands of “novel” structures without clear reporting of which filters have been passed, the signal-to-noise ratio collapses. A standardized hierarchical definition restores clarity, enables fair comparison across models, and aligns computational claims more closely with laboratory realities. The

following sections first document current usage and confusions, then articulate the three core meanings, establish explicit boundary conditions, and finally propose the operational definition that resolves the ambiguity once and for all.

Current Usage and Confusions

The literature reveals four distinct yet overlapping usages of “novel material,” each carrying implicit assumptions that invite overinterpretation. Novelty is frequently framed as absence from the training set, with authors reporting that 95 % of generated crystals lie outside the Materials Project database and treating this metric as conclusive [1, 11, 19, 20]. Such distributional definitions prove computationally convenient yet conceptually shallow: minor perturbations of known prototypes—small shifts in lattice parameters or atomic substitutions within similar radii—satisfy the criterion while remaining firmly within the learned manifold, thereby failing to expand chemical space in any substantive sense.

A related but distinct usage equates novelty with basic chemical validity, whereby structures are deemed novel once they exhibit charge neutrality, plausible bond lengths, and realistic coordination environments [5, 15, 18]. Although necessary for physical plausibility, these elementary checks fall short of establishing meaningful discovery, as charge-balanced configurations may nonetheless prove dynamically unstable or lie far above the convex hull.

This limitation becomes especially acute when novelty is asserted through the absence of prior experimental reports, as in claims of discovering previously unsynthesized superconductors [6, 16]. Here the evidentiary gap risks conflating lack of documentation with genuine synthesizability, overlooking the thermodynamic or kinetic barriers that may render such structures practically inaccessible despite their computational stability.

Beyond these concerns, structural deviation from known prototypes—manifested as new space groups or unprecedented connectivity patterns—is often presented as sufficient innovation [1, 21, 22]. Yet prototype novelty alone guarantees neither chemical validity nor experimental viability, exposing a deeper conceptual slippage wherein single headline figures of “novelty” obscure divergent underlying criteria.

The resulting ambiguity permits systematic overclaims that erode confidence when experimental validation falters. While prior reviews of generative models for inverse design have diagnosed this inconsistency [3-5, 23], they rarely advance an operational remedy; the present framework addresses this gap through mandatory multi-level reporting that disentangles distributional, validity, synthesizability, and structural dimensions of novelty.

Three Meanings of “Novel Material”

Three conceptually distinct meanings emerge from the literature [2, 19].

Table 1 formalizes the ontological distinctions between levels of novelty by linking each definition to its required evidence, operational criteria, and characteristic failure modes.

Table 1. Ontological Separation of Novelty Definitions: Criteria, Evidence Requirements, and Failure Risks Across Hierarchical Levels

Level	Definition Type	Ontological Basis	Operational Criteria	Required Evidence Type	Common Misinterpretation	Failure Risk if Misused
Level 0	Known	Exact identity	Database match	Structural matching	Ignored baseline	False novelty inflation
Level 1	Distributional Novelty	Statistical distance	Descriptor distance >	Dataset comparison +	Treated as discovery	Redundant or trivial structures

			threshold; not in dataset	similarity metric		
Level 2	Chemical Validity	Physical plausibility	Charge neutrality; bond-length constraints; no overlap	Rule-based validation + geometric checks	Treated as synthesizable	Physically plausible but unstable outputs
Level 3	Thermodynamic Stability	Energy minimization	Energy above hull < 0.1 eV/atom	DFT or surrogate energy evaluation	Treated as experimentally viable	Metastable or inaccessible phases
Level 4	Dynamical Stability	Vibrational stability	No imaginary phonon modes	Phonon dispersion analysis	Assumed sufficient for synthesis	Kinetic inaccessibility ignored
Level 5	Experimental Synthesizability	Laboratory feasibility	Realistic synthesis pathway; feasible conditions	Experimental data or pathway simulation	Rarely distinguished explicitly	Misallocation of experimental resources

Distributional novelty designates structures that lie outside the training distribution, operationally defined by non-identity with any training entry, exceeding a chosen similarity threshold in descriptor space, and exclusion from established prototype clusters [1, 11, 19, 20]. In active-learning contexts, this criterion captures genuine exploration of unsampled regions of chemical space, yet it remains silent on whether the generated configuration carries chemical meaning or realizable potential.

Chemical validity imposes a stricter atomic-scale filter, requiring adherence to fundamental physical and chemical constraints: near-zero net charge, bond lengths within accepted ranges, absence of unphysical overlaps, and plausible coordination environments [14, 15, 18]. Some approaches incorporate a loose thermodynamic bound on formation energy to eliminate manifestly unstable compositions. While this layer establishes basic physical plausibility and discards obvious artifacts, it still admits metastable or dynamically unstable candidates that lack deeper viability.

Synthesizability, the most demanding criterion, insists on laboratory accessibility and unfolds across multiple interdependent layers [6, 7, 16, 17]. Thermodynamic stability near or on the convex hull forms the foundational requirement, followed by dynamical stability through the absence of imaginary phonon modes; higher tiers further demand kinetic accessibility via reasonable synthesis conditions and non-hazardous precursors, culminating in experimental realizability aligned with current laboratory capabilities. Only structures that traverse these successive gates offer genuine practical value for experimental follow-up.

These layers stand in irreducible hierarchy: distributional novelty serves as the initial gate, chemical validity the subsequent filter, and synthesizability the culminating test that builds upon both. Papers that bypass intermediate stages inevitably inflate reported novelty metrics. Empirical studies of synthesizability prediction confirm that only a modest fraction of chemically valid candidates ultimately survive thermodynamic and dynamical scrutiny [6, 17]. The present framework therefore mandates explicit reporting of survival rates at each successive gate, enabling readers to assess precisely how many generated structures retain substantive promise.

Boundary Conditions for Each Meaning

Clear thresholds are required if the definitions are to be operational [14].

For distributional novelty the boundary question is “how different is different enough?” The proposal is that a structure counts as novel only if its distance to the nearest training structure exceeds twice the average nearest-neighbor distance inside the training set. When a normalized similarity metric is used, a threshold of 0.3 provides a practical cutoff [1, 19, 22]. Structures falling below this line are better labeled “near-known” rather than novel.

For chemical validity the thresholds are more concrete. Net charge must be smaller than 0.01 electrons per formula unit. Bond lengths must lie within 20 percent of established ranges or two standard deviations of known values for the element pair. Minimum interatomic distance must exceed 0.5 angstroms or 0.7 times the sum of covalent radii. Formation energy should remain below one electron volt per atom to exclude compositions that are obviously unrealistic [14, 18]. These checks can be automated and are already implemented in several generative pipelines [11, 12].

For synthesizability the levels become progressively stricter. Level 1 (thermodynamic stability) requires the energy above hull to be smaller than 0.1 electron volts per atom [6, 16]. Level 2 (dynamical stability) demands no imaginary phonon modes larger than five wavenumbers. Level 3 (kinetic accessibility) adds constraints such as synthesis temperature below 1500 kelvin and avoidance of toxic or explosive precursors. Level 4 (experimental realizability) requires either a documented synthesis route in the literature or a computationally predicted pathway that fits current laboratory capabilities [17].

The hierarchy is strict: a structure cannot reach Level 1 synthesizability without first satisfying chemical validity; Level 2 requires Level 1; and so on. This ordering reflects the logical dependencies in materials evaluation and prevents authors from claiming high-level novelty on the basis of low-level filters alone.

Proposed Operational Definition

A generated material is novel only if it meets explicit criteria at multiple successive levels. The definition is presented as a table of six cumulative levels.

This hierarchical definition forces transparency. It allows readers to see exactly where the model's output stops satisfying increasingly stringent filters. It also aligns computational claims with experimental expectations: only structures that reach Level 5 deserve the unqualified label “novel material” in the strongest sense. By standardizing language across the field, the definition reduces overclaims, improves comparability, and accelerates the pipeline from generative prediction to laboratory validation [8, 19, 24].

Boundary Cases and Gray Zones

Even with a clear hierarchical definition, several boundary cases expose the practical limits of any novelty framework and require explicit discussion.

Table 2 decomposes boundary cases into precise hierarchical failure points, demonstrating how gray-zone structures must be labeled to avoid ambiguity in novelty claims.

Table 2. Analytical Decomposition of Boundary Cases: Mapping Gray-Zone Structures to Hierarchical Failure Points and Reporting Language

Boundary Case	Structural Description	Levels Satisfied	Level of Failure	Correct Reporting Label	Implication for Evaluation	Risk if Misreported

Novel polymorph of known composition	Same composition, different crystal structure	L1, L2	Fails L3+	“Novel structure of known material”	Structural exploration, not discovery	Inflated novelty claims
New composition, known prototype	New elements in familiar topology	L1, L2 (sometimes L3)	Fails L4–L5	“Compositionally novel candidate”	Promising extrapolation region	Premature synthesis claims
Near-threshold similarity	Slightly below novelty cutoff	L0 (borderline L1)	Fails L1	“Near-known structure”	Requires sensitivity analysis	Loss of useful candidates or overclaim
Chemically valid but unstable	Passes bonding rules but high energy	L1, L2	Fails L3	“Chemically valid, unstable candidate”	Needs stability correction	Misleading feasibility
Dynamically stable but kinetically inaccessible	No imaginary modes but extreme synthesis conditions	L1–L4	Fails L5	“Stable but kinetically inaccessible”	Highlights synthesis barriers	Unrealistic experimental expectations

Known compositions can adopt entirely new crystal arrangements, yielding structures that differ in space-group symmetry or atomic coordinates from every training entry and thus qualify as distributionally novel [1, 21, 22]. Such configurations readily satisfy chemical-validity criteria [18], yet they represent novel structures of known materials rather than genuinely new substances. Generative models have already been shown to rediscover established polymorphs while labeling them as novel [3, 4], underscoring the need for precise terminology that distinguishes structural innovation from material discovery.

New compositions constructed on familiar prototypes offer a more promising avenue: the elemental makeup departs from the training distribution while the underlying topology—such as a perovskite framework with an unprecedented A-site occupant—remains validated [7, 11, 16, 20]. These outputs achieve distributional novelty and often pass basic chemical-validity filters when charge balance holds, yet their ultimate value hinges on subsequent thermodynamic and dynamical assessment [17]. In practice, they embody genuine compositional expansion within structurally reliable families [6].

Boundary cases further complicate assessment. Structures falling just inside a chosen distributional threshold—perhaps registering a similarity score marginally below the cutoff—technically reside within the training manifold yet may still exhibit distinct properties. Transparent reporting as “near-known” together with the precise distance metric avoids both overclaiming and premature dismissal of potentially useful candidates [19, 23].

Chemically valid configurations that lie appreciably above the convex hull—such as 0.2 eV per atom—likewise demand nuanced classification: they obey local bonding rules but fail the foundational thermodynamic gate and therefore cannot be considered synthesizable at the primary level. Under specific conditions, however, they may persist as metastable phases [6, 17], warranting description as “chemically valid, metastable candidates” rather than novel materials.

Even dynamically stable structures can encounter insurmountable practical barriers when required synthesis conditions exceed accessible laboratory regimes or involve hazardous precursors, satisfying the dynamical criterion while failing higher kinetic and experimental layers [7, 16]. Explicit annotation of each attained level is therefore essential, allowing experimentalists to gauge realistic barriers without ambiguity [6, 7, 17].

These boundary zones reveal why isolated novelty percentages remain inadequate. Only a hierarchical framework that annotates every generated structure according to the precise level it achieves can eliminate interpretive slippage and

effectively guide downstream experimental validation efforts [4, 5].

Relation to Other Concepts

The proposed definition of “novel material” sits within a broader ecosystem of concepts already discussed in the generative-model literature [15, 24–26].

In generative chemistry for molecules, novelty is more maturely defined through metrics such as Tanimoto similarity on molecular fingerprints or scaffold novelty [5, 14, 15, 24]. These measures operate in a finite, non-periodic chemical space where “different enough” is easier to quantify. Crystal novelty, by contrast, must contend with periodicity, variable unit-cell sizes, and coupled composition-structure constraints, making direct transfer of molecular metrics inappropriate [4, 9, 27].

Chemical-validity metrics reviewed in recent benchmarking studies already capture charge balance, bond-length plausibility, and coordination rules [14, 18]. The present framework builds directly on those metrics but treats them as an intermediate gate rather than the final claim. Chemical validity is necessary yet insufficient for declaring a material novel; synthesizability remains the decisive filter.

Stability prediction appears throughout the literature as an isolated task [6, 7, 16]. Convex-hull distance and phonon analysis are routinely applied after generation, yet authors frequently present stable structures as “novel” without acknowledging that stability is only one layer within the broader synthesizability hierarchy. The operational definition integrates these stability checks explicitly, clarifying their position relative to distributional and validity filters.

Extrapolation is another closely related idea [1, 28]. Distributional novelty is precisely a measure of extrapolation beyond the training manifold. Earlier work on generative models for inorganic materials already notes the risk of mode collapse and limited extrapolation [3, 4]. The hierarchical definition adds chemical-validity and synthesizability layers on top of extrapolation, transforming a purely statistical concept into a chemically and practically meaningful one.

By situating the new definition at the intersection of these established concepts, the framework avoids reinventing terminology while supplying the missing boundary conditions that previous reviews identified but did not formalize [5, 14, 15].

Implications for Generative Model Evaluation

The hierarchical definition carries direct consequences for three audiences: model developers, benchmark designers, and experimental practitioners.

For model developers the requirement is unambiguous. Every paper must report novelty percentages at each level rather than a monolithic “X % novel” statistic [3, 21, 23]. Developers should state the exact similarity threshold used for distributional novelty, the precise charge-balance and bond-length tolerances applied, and the energy-above-hull cutoff for thermodynamic stability. Vague claims such as “our model discovers novel materials” become unacceptable once the community adopts the standard [3–5].

Benchmark designers gain a ready-made evaluation rubric [21, 23]. Future generative-model benchmarks should include the full hierarchy as mandatory metrics: percentage distributionally novel, percentage chemically valid, percentage thermodynamically stable, percentage dynamically stable, and percentage potentially synthesizable. Leaderboards that currently rank models on single validity scores or hull distances can be extended to track progression through all six levels, revealing which models truly push the frontier of actionable materials [14].

For practitioners—both computational and experimental—the framework supplies a transparent screening funnel. Experimental groups can focus resources on the small subset of outputs that reach Level 5 while discarding the much

larger fraction that fail earlier gates. Computational workflows can embed the hierarchical checks directly into generation pipelines, rejecting invalid candidates early and reducing downstream computational waste [6, 7, 10, 16].

Collectively these implications shift evaluation culture from headline novelty numbers to traceable, multi-stage filtering. The result is fairer model comparison, reduced overclaim, and a tighter coupling between computational output and laboratory feasibility [17, 29].

Toward a Standardized Novelty Reporting Protocol

To operationalize the hierarchical definition, a standardized reporting format should become mandatory in every generative-model paper. Authors would present the multi-level results in a compact, prose-style summary immediately after the results section. A typical report would read as follows: distributional novelty reached 95 % (structures not in training set with distance greater than 0.3); chemical validity reached 80 % charge-balanced and 85 % with reasonable bond lengths; thermodynamic stability reached 40 % with energy above hull below 0.1 eV/atom; dynamic stability reached 20 % with no imaginary phonons; and overall synthesizability reached 10 % meeting Levels 1–3.

Interpretation follows automatically: “Our model generates 95 % distributionally novel structures. After filtering for chemical validity and stability, 10 % of generated structures are potentially synthesizable.”

Such a standardized summary eliminates ambiguity at a glance. Reviewers and readers can immediately assess the practical value of the generated candidates rather than deciphering prose claims scattered throughout the text. The protocol also encourages authors to document the exact thresholds and software used for each filter, enhancing reproducibility [14, 15, 18].

Community adoption requires only a modest cultural shift. Journals in computational materials science—already accustomed to demanding open-source code and deposited structures—can add a novelty-reporting summary to their checklist. Conferences and workshops can feature the required format in example submission templates. Once a critical mass of papers follows the protocol, non-compliant claims will stand out as incomplete [3, 4, 23].

The payoff is substantial. Standardized multi-level reporting will reduce the signal-to-noise ratio in the literature, accelerate meta-analyses across models, and build credibility with experimental collaborators who have grown skeptical of unfiltered generative claims [5, 17, 26]. The framework is ready for immediate use; the only remaining step is collective agreement to apply it.

Conclusion

“Novel material” is an ambiguous term that has been used inconsistently across the generative-model literature. Three distinct meanings—distributional novelty (not in the training set), chemical validity (obeys physical and chemical laws), and synthesizability (can be made in a laboratory)—are routinely conflated, producing overclaims that hinder progress. This boundary/definitional article has mapped the conceptual boundaries, established explicit thresholds for each meaning, and proposed a six-level hierarchical operational definition: Known → Distributionally novel → Chemically valid → Thermodynamically stable → Dynamically stable → Experimentally synthesizable.

Boundary cases and gray zones demonstrate that the hierarchy forces precise language rather than blanket assertions. The definition aligns with related concepts in molecular generative chemistry, stability prediction, and extrapolation while adding the missing synthesizability dimension. Its implications extend to model evaluation, benchmark design, and practical screening workflows.

The call to action is straightforward: every generative materials paper should include the standardized novelty-reporting table and specify the level(s) attained by its claims. Adoption of this protocol will replace rhetorical novelty with traceable,

multi-stage filtering, restore trust between computational and experimental communities, and accelerate the discovery of materials that are not merely generated but genuinely synthesizable. The boundary between chemical validity and synthesizability is now clearly drawn; the field can move forward with a shared, operational language.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 22 Jul 2024

Revised: 01 Oct 2024

Accepted: 24 Dec 2024

Published online: 18 January 2025

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

- Zeni C, Pinsler R, Zügner D, Fowler A, Horton M, Fu X, et al. A generative model for inorganic materials design. *Nature*. 2025;639(8055):624-32.
<https://doi.org/10.1038/s41586-025-08628-5>.
- Dong R, Fu N, Siriwardane EM, Hu J. Generative design of inorganic compounds using deep diffusion language models. *J Phys Chem A*. 2024;128(29):5980-9.
<https://doi.org/10.1021/acs.jpca.4c00083>.
- Chen Z, Meng Z, He T, Li H, Cao J, Xu L, et al. Crystal structure prediction meets artificial intelligence. *J Phys Chem Lett*. 2025;16(10):2581-91.
<https://doi.org/10.1021/acs.jpclett.4c03727>.
- De Breuck PP, Wang HC, Rignanese GM, Botti S, Marques MAL. Generative AI for crystal structures: A review. *npj Comput Mater*. 2025;11(1):370.

<https://doi.org/10.1038/s41524-025-01881-2>.

Anstine DM, Isayev O. Generative models as an emerging paradigm in the chemical sciences. *J Am Chem Soc.* 2023;145(16):8736-50.

<https://doi.org/10.1021/jacs.2c13467>.

Antoniuk ER, Cheon G, Wang G, Bernstein D, Cai W, Reed EJ. Predicting the synthesizability of crystalline inorganic materials from the data of known material compositions. *npj Comput Mater.* 2023;9(1):155.

<https://doi.org/10.1038/s41524-023-01114-4>.

Davariashtiyani A, Kadhodaie Z, Kadhodaie S. Predicting synthesizability of crystalline materials via deep learning. *Commun Mater.* 2021;2(1):115.

<https://doi.org/10.1038/s43246-021-00219-x>.

Ardizzone L, Kruse J, Wirkert S, Rahner D, Pellegrini EW, Klessen RS, et al. Analyzing inverse problems with invertible neural networks. *arXiv [Preprint]*. 2018.

<https://doi.org/10.48550/arXiv.1808.04730>.

Wan XH, Zhang Y, Guo QH, Zheng LY. Deep learning-based inverse design of irregular phononic crystals. *Int J Mech Sci.* 2025;297-298:110335.

<https://doi.org/10.1016/j.ijmecsci.2025.110335>.

Banik S, Loeffler T, Manna S, Chan H, Srinivasan S, Darancet P, et al. A Continuous Action Space Tree search for INverse desiGn (CASTING) framework for materials discovery. *npj Comput Mater.* 2023;9(1):177.

<https://doi.org/10.1038/s41524-023-01128-y>.

Dan Y, Zhao Y, Li X, Li S, Hu M, Hu J. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. *npj Comput Mater.* 2020;6(1):84.

<https://doi.org/10.1038/s41524-020-00352-0>.

Manica M, Born J, Cadow J, Christofidellis D, Dave A, Clarke D, et al. Accelerating material design with the generative toolkit for scientific discovery. *npj Comput Mater.* 2023;9(1):69.

<https://doi.org/10.1038/s41524-023-01028-1>.

Karpovich C, Pan E, Olivetti EA. Deep reinforcement learning for inverse inorganic materials design. *npj Comput Mater.* 2024;10(1):287.

<https://doi.org/10.1038/s41524-024-01474-5>.

Gao W, Coley CW. The synthesizability of molecules proposed by generative models. *J Chem Inf Model.* 2020;60(12):5714-23.

<https://doi.org/10.1021/acs.jcim.0c00174>.

Chen Z, Min MR, Parthasarathy S, Ning X. A deep generative model for molecule optimization via one fragment modification. *Nat Mach Intell.* 2021;3(12):1040-9.

<https://doi.org/10.1038/s42256-021-00410-2>.

Zhu R, Tian SIP, Ren Z, Li J, Buonassisi T, Hippalgaonkar K. Predicting synthesizability using machine learning on databases of existing inorganic materials. *ACS Omega.* 2023;8(9):8210-8.

<https://doi.org/10.1021/acsomega.2c04856>.

Chung V, Walsh A, Payne DJ. Solid-state synthesizability predictions using positive-unlabeled learning from human-curated literature data. *Digit Discov.* 2025;4(9):2439-53.

<https://doi.org/10.1039/D5DD00065C>.

Gleaves D, Fu N, Siriwardane EM, Zhao Y, Hu J. Materials synthesizability and stability prediction using a semi-supervised teacher-student dual neural network. *Digit Discov.* 2023;2(2):377-91.

<https://doi.org/10.1039/D2DD00098A>.

Luo X, Wang Z, Wang Q, Shao X, Lv J, Wang L, et al. CrystalFlow: A flow-based generative model for crystalline materials. *Nat Commun.* 2025;16(1):9267.

<https://doi.org/10.1038/s41467-025-64364-4>.

Zhou Z, Shang Y, Liu X, Yang Y. A generative deep learning framework for inverse design of compositionally complex bulk metallic glasses. *npj Comput Mater.* 2023;9(1):15.

<https://doi.org/10.1038/s41524-023-00968-y>.

Barra J, Chahal R, Banerjee S, Lupo Pasini M, Irle S, Lam S. Inverse mapping of properties to composition through generative modeling for designing molten salts. *npj Comput Mater.* 2025;11(1):190.

<https://doi.org/10.1038/s41524-025-01638-x>.

Fung V, Zhang J, Hu G, Ganesh P, Sumpster BG. Inverse design of two-dimensional materials with invertible neural networks. *npj Comput Mater.* 2021;7(1):200.

<https://doi.org/10.1038/s41524-021-00670-x>.

Lyngby P, Thygesen KS. Data-driven discovery of 2D materials by deep generative models. *npj Comput Mater.* 2022;8(1):232.

<https://doi.org/10.1038/s41524-022-00923-3>.

Kim K, Kang S, Yoo J, Kwon Y, Nam Y, Lee D, et al. Deep-learning-based inverse design model for intelligent discovery of organic molecules. *npj Comput Mater.* 2018;4(1):67.

<https://doi.org/10.1038/s41524-018-0128-1>.

Qiu H, Sun ZY. On-demand reverse design of polymers with PolyTAO. *npj Comput Mater.* 2024;10(1):273.

<https://doi.org/10.1038/s41524-024-01466-5>.

Kumar JN, Li Q, Tang KYT, Buonassisi T, Gonzalez-Oyarce AL, Ye J. Machine learning enables polymer cloud-point engineering via inverse design. *npj Comput Mater.* 2019;5(1):73.

<https://doi.org/10.1038/s41524-019-0209-9>.

Skinnider MA, Wang F, Pasin D, Greiner R, Foster LJ, Dalsgaard PW, et al. A deep generative model enables automated structure elucidation of novel psychoactive substances. *Nat Mach Intell.* 2021;3(11):973-84.

<https://doi.org/10.1038/s42256-021-00407-x>.

Park H, Onwuli A, Walsh A. Exploration of crystal chemical space using text-guided generative artificial intelligence. *Nat Commun.* 2025;16(1):4379.

<https://doi.org/10.1038/s41467-025-59636-y>.

Guo L, Ren Y, Li R, Jiang B. Prior knowledge-augmented broad reinforcement learning framework for fault diagnosis of heterogeneous multiagent systems. *IEEE Trans Cogn Dev Syst.* 2024;16(1):115-23.

<https://doi.org/10.1109/TCDS.2023.3266791>.