O'Connor *et al. Journal of Artificial Intelligence for Materials Science* *(2026) 1:148*
https://doi.org/10.51847/c392803226

ORIGINAL RESEARCH **Open access**

The Problem of Scientific Path Abandonment in AI-Guided Materials Research

Patrick O'Connor[1*], Sean Murphy[1]

Abstract

Scientific path abandonment has emerged as a critical yet underrecognized failure mode in AI-guided materials research, in which promising research directions—such as novel compositional families, structural motifs, or synthesis routes—are terminated prematurely due to insufficient evidence, narrow optimization signals, or algorithmic impatience. This failure mode is defined as the termination of a research direction before sufficient evidence has been gathered to determine its true promise, distinguishing it from rational stopping grounded in conclusive data. The mechanisms driving this abandonment include algorithmic impatience that halts exploration upon short-term metric plateaus, overconfidence in negative predictions, narrow optimization that sacrifices multi-objective potential, and exploration decay inherent in active learning loops. Four distinct types of path abandonment—compositional, structural, synthesis, and property—each generate specific failure modes, such as local optima traps, false-negative cascades, exploration starvation, and regret amplification. Detection principles center on systematic audits, counterfactual reasoning, diversity monitoring, and regret tracking. In contrast, mitigation principles emphasize extended exploration, resource reserves, delayed abandonment thresholds, path revisitation, and regret-aware stopping rules. By articulating this failure mode and offering a comprehensive framework for recognition and remedy, the analysis identifies scientific path abandonment as a systemic risk that undermines the very autonomy and discovery potential that AI promises to deliver in materials science.

Keywords Scientific path abandonment, AI-guided materials research, Premature termination, Active learning convergence, Opportunity cost in discovery, Regret-aware optimization

*Correspondence:
Patrick O'Connor
patrick.oconnor@gmail.com
[1] Department of Materials Informatics and AI Systems, Trinity College Dublin, Dublin, Ireland

Introduction

In the rapidly evolving landscape of AI-guided materials research, autonomous systems now routinely make high-stakes decisions about which research paths to pursue and which to abandon. These decisions, while intended to accelerate discovery, often lead to the premature termination of promising avenues, thereby eroding the overall efficiency and creativity of the materials discovery pipeline. The problem is not merely one of computational efficiency but of epistemic integrity: AI systems, operating under tight resource constraints and optimization pressures, often terminate exploration of compositional spaces, structural variants, or synthesis protocols before sufficient data have accumulated to reveal their latent value. This paper analyzes scientific path abandonment as a distinct failure mode in AI-driven materials science, arising at the intersection of algorithmic decision rules, human cognitive biases inherited by surrogate models, and the inherent uncertainty of high-dimensional materials spaces [1-3].

The consequences of such premature abandonment extend far beyond individual projects. When an AI agent discards an entire family of halide perovskites after observing modest bandgap performance in the first few candidates, or abandons a synthesis route for solid-state electrolytes after two unsuccessful attempts, the opportunity costs accumulate silently. These costs manifest as delayed breakthroughs in energy storage, catalysis, or quantum materials, as well as the wasteful rediscovery of previously abandoned paths by later researchers. Existing literature on machine learning for materials has predominantly focused on model accuracy, data efficiency, and autonomous experimentation platforms, yet has paid comparatively little attention to the stopping rules and abandonment logics embedded in these systems [4-7].

This failure mode analysis, therefore, identifies scientific path abandonment as the premature termination of research directions driven by AI decision rules, optimization pressures, or inherited cognitive biases. It builds directly upon foundational work that highlights the challenges of autonomous materials research, noting that while AI promises to navigate vast design spaces, it simultaneously risks truncating exploration in ways that rational human investigators might avoid [7]. The analysis proceeds by first establishing a precise conceptual definition, then dissecting the underlying mechanisms, cataloging the primary types of abandonment encountered in materials contexts, and finally delineating the specific failure modes that result. Throughout, the discussion remains grounded in the conceptual and decision-theoretic dimensions of the problem rather than empirical performance metrics. By the conclusion of this first part, the framework will have articulated a typology that equips researchers to recognize and interrogate abandonment decisions within their own workflows. The ultimate goal is not to eliminate abandonment—since rational pruning remains essential—but to ensure that abandonment occurs only when evidence, rather than algorithmic impatience, justifies termination [3, 8-14].

The broader context of this failure mode is the increasing reliance on active learning, Bayesian optimization, and reinforcement learning frameworks that embed implicit stopping criteria. These frameworks, while powerful, inherit the classic "stopping problem" from statistical decision theory, where the optimal moment to cease experimentation is theoretically intractable without perfect foresight. In materials science, the stakes are particularly high because the design space is combinatorially explosive and the payoff landscape is rugged and multi-modal. An AI system that abandons paths too readily may converge on local optima while missing transformative global discoveries; conversely, one that never abandons risks resource exhaustion. The present analysis, therefore, situates path abandonment as a failure mode that is both epistemic—concerning what we fail to learn—and practical—concerning the allocation of scarce experimental and computational resources [15-18].

Defining Path Abandonment

Scientific path abandonment must be defined with precision if it is to serve as a diagnostic category within AI-guided materials research. Definition 1: Scientific Path Abandonment is the termination of a research direction—whether a specific compositional family, structural motif, synthesis method, or property target—before sufficient evidence has been gathered to determine its true promise, thereby foreclosing further investigation that could have revealed substantial value. This definition emphasizes the temporal and evidentiary insufficiency of the termination decision rather than the absolute absence of progress.

It is essential to distinguish scientific path abandonment from three related but conceptually distinct processes. First, rational stopping occurs when accumulated evidence reaches a statistically defensible threshold that conclusively demonstrates the direction's limited promise; here, termination is epistemically justified rather than premature. Second, pruning refers to the elimination of demonstrably dominated options—those that are provably inferior across all relevant objectives based on already available data—without discarding potentially non-dominated alternatives. Third, switching involves the reallocation of resources from one direction to another without necessarily declaring the first direction unpromising; it is a tactical re-prioritization rather than a permanent epistemic closure. Path abandonment, by contrast, enacts a stronger form of foreclosure: the direction is removed from the active consideration set and often from institutional memory, rendering later revisitation costly or impossible [3, 7].

Table 1 differentiates scientific path abandonment from rational stopping, pruning, and switching by clarifying how evidentiary sufficiency, scope of closure, and reversibility diverge across termination logics.

Table 1. Conceptual demarcation of scientific path abandonment from adjacent termination logics in AI-guided materials research

Category	Core definition	Evidentiary status at termination	Scope of action
Rational stopping	Termination after sufficient evidence demonstrates limited promise	Evidence is judged adequate relative to the decision threshold	Stops a path judged conclusively weak
Pruning	Elimination of clearly dominated options while preserving non-dominated alternatives	Evidence is sufficient for local comparative elimination, not whole-path closure	Removes subsets, no necessarily entire directions
Switching	Resource reallocation from one path to another without a final judgment of worthlessness	Evidence may be incomplete, but termination is tactical rather than final	Suspends emphasis rather than closes the path
Scientific path abandonment	Termination of a research direction before sufficient evidence has been gathered to determine true promise	Evidence is insufficient relative to the magnitude and uncertainty of the path	Removes th path from th active consideratio set, often from memor as well

The distinction carries practical weight in materials contexts. Consider, for instance, an AI-orchestrated campaign targeting high-entropy oxides. A rational stopping rule might terminate after exhaustive sampling confirms thermodynamic instability across the entire family. Pruning might discard only those compositions whose predicted formation energies exceed a clear threshold while retaining borderline candidates. Switching might deprioritize the oxide family temporarily to allocate budget to a parallel sulfide campaign. Abandonment, however, would strike the entire family from the search graph after observing only a handful of mediocre candidates, even though latent stable phases or unexpected electronic properties might exist deeper in the space. This last process is what the present framework identifies as the failure mode.

A conceptual diagram of the abandonment decision process can be visualized as a two-dimensional trade-off surface. The horizontal axis represents cumulative evidence (experimental or simulated data points), while the vertical axis represents the estimated promise of the path (expected utility or multi-objective score). A dynamic threshold curve slopes upward with increasing evidence: early in the process, the threshold is low, permitting continuation despite uncertainty; as evidence accumulates, the threshold rises, reflecting the growing cost of continued investment. AI systems frequently place or cross this threshold too aggressively—either by inflating the slope through overconfident surrogate models or by applying a static high threshold derived from narrow single-objective optimization—thereby triggering abandonment well before the true promise curve (which may rise sharply later) justifies it. The gap between the algorithmic threshold and the counterfactual true-promise trajectory constitutes the visual signature of premature path abandonment.

Mechanisms of Abandonment

Four interlocking mechanisms drive scientific path abandonment in AI-guided materials research. Each operates through distinct algorithmic and decision-theoretic pathways yet converges on the same outcome: premature termination.

Impatience

AI systems, particularly those employing Bayesian optimization or reinforcement learning, often terminate paths when short-term performance metrics plateau, even though longer-horizon gains remain probable. This impatience arises because surrogate models are trained to maximize immediate expected improvement, discounting distant rewards to save computational resources. In materials discovery, where property landscapes are noisy and multi-scale, a temporary plateau in predicted conductivity or stability does not preclude later breakthroughs once additional descriptors or higher-fidelity simulations are incorporated. Yet the algorithm interprets the plateau as diminishing returns and abandons the path [18,19-28].

Overconfidence

Modern surrogate models often assign excessively narrow uncertainty estimates to negative predictions, leading to high-confidence rejections of paths that are, in reality, only marginally suboptimal. This overconfidence is an artifact of Gaussian process assumptions or deep-ensemble calibration failures in high-dimensional spaces. When an AI agent predicts with 95% posterior probability that a given perovskite composition will never achieve a target bandgap, it abandons the composition despite the model's acknowledged epistemic uncertainty. The mechanism is particularly insidious because it masquerades as rigorous probabilistic reasoning while masking the model's own epistemic debt [1, 21].

Narrow optimization

When AI frameworks optimize a single scalar objective—band gap, formation energy, or ionic conductivity—they implicitly abandon paths that are mediocre on that objective but exceptional on unmodeled or secondary objectives. Multi-objective optimization can mitigate this, yet even Pareto-front approaches often collapse to a weighted sum that privileges one axis. A synthesis route abandoned because it yields slightly lower throughput may nonetheless enable previously inaccessible metastable phases; the narrow optimizer never registers this latent value [4, 7].

Exploration decay

Active learning inherently reduces exploration as the surrogate model's confidence grows. Query strategies shift from uncertainty sampling to exploitation, causing the algorithm to sample ever more densely around current optima while progressively ignoring distant, unexplored regions of composition or structure space. Over time, entire manifolds of the design space are de facto abandoned, not because they were evaluated and found wanting, but because they were never evaluated at all. This decay is mathematically guaranteed in finite-budget settings and is exacerbated by budget constraints typical of experimental materials campaigns [21, 22].

These mechanisms rarely operate in isolation. Impatience amplifies overconfidence, narrow optimization accelerates exploration decay, and the resulting feedback loop can terminate dozens of paths within a single campaign. The present analysis articulates these mechanisms to render them visible and therefore contestable within workflow design.

Types of Path Abandonment

Path abandonment in materials research manifests in four primary types, each corresponding to a distinct dimension of the materials discovery space.

Compositional abandonment

This occurs when an AI system terminates exploration of an entire compositional family—e.g., all A-site doped perovskites within a given tolerance factor range—after sampling only a sparse subset. Definitely, the family is removed from the search graph without exhaustive coverage. In practice, this might manifest as the early rejection of all lead-free double perovskites after observing poor predicted stability in the first 10 candidates, foreclosing potential high-performance variants that lie farther into the chemical space. The consequence is the permanent loss of a broad region of the periodic table design space [5, 6].

Structural abandonment

Here, the AI abandons specific structural motifs—layered Ruddlesden-Popper phases, spinel derivatives, or metal-organic framework topologies—before fully enumerating their compositional variants. A common scenario involves the early dismissal of 2D layered structures after modest predicted exfoliation energies, even though subtle cation ordering or vacancy engineering could unlock exotic electronic or magnetic properties. The type is particularly damaging because structural motifs often encode emergent phenomena that compositional tweaks merely modulate.

Synthesis abandonment

This type concerns the premature termination of synthesis routes—solid-state, solvothermal, or vapor-phase—after a handful of experimental or simulated failures. An AI-orchestrated robotic lab might abandon a flux-growth protocol for a new cathode material after two unsuccessful runs, ignoring the possibility that minor adjustments in temperature ramp or precursor stoichiometry could succeed. The consequence is the loss of kinetically accessible metastable phases that thermodynamic databases cannot predict [7].

Property abandonment

The AI declares a target property—room-temperature superconductivity, ultrahigh piezoelectric coefficient, or negative thermal expansion—“unreachable” and abandons all paths aimed at it. This occurs when surrogate models extrapolate pessimistically beyond the training manifold, effectively erasing an entire objective from the research agenda. The type carries the highest strategic regret because property targets often define grand challenges in energy, quantum information, or sustainability.

Each type shares the core signature of evidentiary insufficiency, yet differs in the granularity of foreclosure and the downstream scientific cost.

A Typology of Abandonment Failure Modes

The four types of path abandonment give rise to four corresponding failure modes that undermine the integrity of AI-guided discovery.

Local optima trap

This failure mode arises when abandonment decisions systematically eliminate paths that initially appear inferior yet serve as bridges to superior global optima. The mechanism is impatience coupled with narrow optimization: the algorithm abandons “downhill” compositions or structures because they degrade the current objective, never discovering that continued exploration would reveal a saddle point leading to a new basin. In materials terms, this might occur when an AI discards a high-entropy alloy series that temporarily lowers ductility but later enables unprecedented corrosion resistance through nanoscale ordering. Detection signature: abrupt shrinkage of the explored volume in composition space accompanied by stagnation of the Pareto front [1, 22].

False negative cascade

Early negative results, amplified by overconfident surrogates, trigger a cascade of further abandonments that propagates across related paths. A single low-fidelity simulation reporting poor stability can cause the AI to prune dozens of neighboring compositions, even though higher-fidelity validation would have overturned the initial prediction. The mode is self-reinforcing: the surrogate is retrained on an increasingly biased dataset that contains only “confirmed” negatives, deepening the false-negative bias. Materials example: an early DFT misprediction of thermodynamic instability in a family of transition-metal dichalcogenides led to wholesale abandonment, only to be later contradicted by experimental synthesis [21].

Exploration starvation

Resources become concentrated on a shrinking set of high-confidence paths while the remainder of the design space is starved of queries. The mechanism is exploration decay accelerated by budget limits. The signature is a rapid collapse in the diversity of sampled candidates, as measured by a drop in the average pairwise distance within the feature space of active queries. In practice, this starves the discovery of outlier materials—those lying far from known clusters but possessing disruptive properties [22].

Regret amplification

Abandoned paths are later demonstrated—by independent research or serendipitous re-exploration—to possess high value, yet the original campaign cannot restart them without prohibitive cost. Regret is amplified because institutional knowledge of the abandonment decision fades, and the opportunity cost is never logged. The mode is the ultimate epistemic penalty of path abandonment: not only is value lost, but the loss itself becomes invisible [3, 28, 29].

Figure 1 maps the hierarchical causal architecture through which AI decision rules prematurely terminate materials research paths, converting local stopping judgments into systematic epistemic foreclosure.

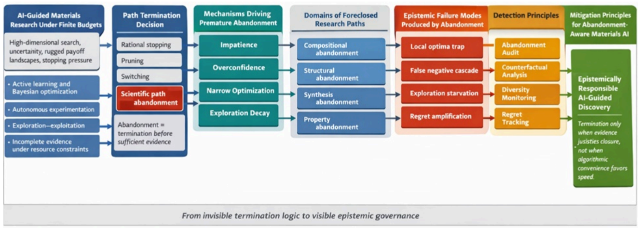

Figure 1. Hierarchical architecture of scientific path abandonment in AI-guided materials research.

These modes are not mutually exclusive; a single campaign may exhibit all four simultaneously. The typology offered here provides a diagnostic vocabulary for interrogating AI workflows at the point of abandonment decisions.

Detection Principles

Detecting scientific path abandonment before it compounds into irreversible opportunity costs requires a set of principled diagnostic practices that operate at the level of workflow metadata rather than post-hoc performance. The present framework proposes four interlocking detection principles that can be embedded directly into AI-guided materials research pipelines.

Abandonment audit

Every termination decision must be logged with explicit metadata capturing the exact moment of abandonment, the surrogate-model confidence at that point, the number of samples evaluated, and the dominant mechanism (impatience, overconfidence, narrow optimization, or exploration decay) inferred from the stopping rule. This audit trail transforms abandonment from an invisible background process into a visible epistemic event. In practice, the audit would flag a compositional family abandoned after only 12 queries, despite an uncertainty band still spanning 0.8 eV in the predicted bandgap, thereby identifying potential prematurity. The principle draws directly on established stopping-criterion literature to ensure that the audit itself remains computationally lightweight yet epistemically transparent [18, 28].

Counterfactual analysis

For any abandoned path, the framework proposes a "what-if" continuation simulation in which the same computational budget is reallocated to that path alone, using the final surrogate model to extrapolate likely outcomes under prolonged exploration. While not claiming predictive certainty, this counterfactual exercise quantifies the plausible regret gap between the abandoned trajectory and the observed campaign optimum. When the extrapolated promise curve crosses the abandonment threshold only after additional virtual samples, the detection principle identifies the original decision as premature. Counterfactual analysis is particularly potent for distinguishing rational stopping from abandonment because it forces explicit comparison against the evidentiary threshold articulated in Definition 1 [3, 21].

Diversity monitoring

The framework identifies a rapid collapse in the diversity metric of the active query set—measured, for example, as average pairwise distance in a learned latent space of compositions or structures—as a reliable signature of exploration starvation and impending path abandonment. When diversity drops below a dynamically adjusted baseline derived from the initial exploration phase, the system triggers an alert that entire manifolds are being de facto abandoned without explicit termination. This principle is especially relevant in materials contexts where structural motifs or synthesis routes may lie in sparsely sampled regions; monitoring diversity therefore serves as an early-warning system before formal abandonment is declared [22, 24]. Regret Tracking. Rather than waiting for external rediscovery of abandoned paths, the framework proposes continuous internal regret tracking that periodically re-evaluates a small reservoir of previously terminated directions using any newly acquired external data or higher-fidelity oracles. A spike in predicted utility for an archived path constitutes the detection signature of regret amplification. By maintaining this live regret ledger, the detection principle converts latent epistemic debt into an actionable signal that can prompt path revisitation before institutional memory fades. Together, these four principles form a detection layer that operates continuously rather than retrospectively, ensuring that scientific path abandonment is recognized at the moment it threatens to become a failure mode, rather than after irreversible foreclosure [3, 28].

The detection principles are not intended as rigid algorithmic add-ons but as reflexive practices that encourage researchers to interrogate abandonment decisions in real time. When applied consistently, they shift the epistemic culture of AI-guided materials research from treating termination as an unexamined optimization byproduct to treating it as a decision requiring explicit justification.

Mitigation Principles

Preventing premature scientific path abandonment requires proactive interventions that alter decision rules, resource-allocation logic, and cultural norms surrounding termination. The framework offers five concrete mitigation principles that can be implemented without sacrificing the core efficiency gains of AI-driven discovery.

Extended exploration

The framework proposes that active-learning loops incorporate a mandatory “extension buffer” that continues sampling along any path for a fixed number of additional iterations after the nominal stopping criterion is met, provided the uncertainty band continues to overlap with the target property region. This buffer directly counters impatience and exploration decay by ensuring that short-term plateaus do not trigger foreclosure immediately. In compositional or structural campaigns, the extension buffer might translate to twenty extra queries per family, a modest computational overhead that nevertheless preserves access to latent promise [18, 22].

Exploration reserve

A fixed percentage of the total experimental or computational budget—typically 15%–20%—is ring-fenced as an “exploration reserve” that can only be spent on paths previously marked for abandonment or on entirely novel regions flagged by diversity monitoring. This reserve serves as a deliberate counterweight to exploitation bias, ensuring that resources remain available for paths that narrow-optimization would otherwise eliminate. The principle recognizes that the true cost of abandonment is not merely the immediate query budget but the foreclosed future value; the reserve therefore functions as an insurance policy against regret amplification [21, 28].

Delayed abandonment

Rather than permitting termination based on a single negative result or a narrow confidence threshold, the framework requires a multi-stage confirmation protocol: at least three independent negative evaluations, spaced across different surrogate-model retraining epochs, before a path may be formally abandoned. This delay directly mitigates overconfidence and false-negative cascades by repeatedly forcing the system to confront its own epistemic uncertainty. In synthesis-route campaigns, for example, a flux-growth protocol would receive three distinct temperature-stoichiometry perturbations before being declared unviable [7, 19].

Path revisitation

The framework articulates a scheduled revisitation protocol in which a random or diversity-weighted subset of abandoned paths is reintroduced into the active set at fixed intervals (every 10% of the total campaign budget). Revisitation is guided by the latest surrogate model augmented with any external data acquired since abandonment. This principle transforms abandonment from a permanent epistemic closure into a reversible state, directly addressing the institutional memory loss that otherwise compounds regret [3, 24].

Regret-aware stopping

Traditional stopping rules are augmented with an explicit regret term that penalizes termination decisions proportionally to the estimated future value of the abandoned path. The regret-aware objective, therefore, balances immediate expected improvement against the long-term cost of foreclosure, producing stopping thresholds that are dynamically sensitive to the multi-objective and rugged nature of materials landscapes. When integrated into Bayesian optimization or reinforcement-learning controllers, this principle ensures that the stopping problem is solved not merely for efficiency but for epistemic completeness [28].

Table 2 consolidates the paper's central theoretical architecture by showing how each abandonment mechanism selectively forecloses certain research paths, produces characteristic failure modes, and demands a corresponding detection or mitigation response.

Table 2. Theoretical consolidation matrix linking abandonment mechanisms, foreclosed path types, failure modes, detection signals, and mitigation levers

Mechanism	Primary form of epistemic distortion	Path types are most vulnerable	Typical failure mode generate
Impatience	Short-horizon plateau is misread as a terminal lack of promise	Compositional abandonment; synthesis abandonment	Local optima tra

Overconfidence	Negative predictions are treated as more certain than the evidence warrants	Compositional abandonment; property abandonment	False negative cascade
Narrow optimization	Single-objective success criteria erase latent multi-objective value	Structural abandonment; synthesis abandonment; property abandonment	Local optima tra regret amplificatio
Exploration decay	Search diversity contracts until entire regions are ignored without explicit justification	Structural abandonment; compositional abandonment	Exploratio starvatio
Mechanism interaction	Multiple distortions combine and reinforce each other	All four path types	All four failure modes ca co-occur

Collectively, these mitigation principles do not eliminate abandonment—an essential component of any finite-resource campaign—but they raise the evidentiary bar and lower the threshold for reversibility, thereby aligning AI decision rules more closely with the long-horizon, multi-objective character of genuine materials discovery. Implementation requires only modest changes to existing open-source active-learning libraries yet yields substantial protection against the failure modes articulated earlier [4, 7].

Relation to Other Failure Modes

Scientific path abandonment does not exist in isolation; it interacts dynamically with several other recognized failure modes in AI-guided science, amplifying their severity while simultaneously being amplified by them. The framework articulates these relations to situate path abandonment within the broader failure-mode landscape of autonomous materials research.

First, path abandonment is closely linked to path dependence. Once a direction is abandoned, the search graph is permanently pruned, foreclosing not only the immediate path but all downstream branches that might have emerged from continued exploration. This foreclosure creates irreversible path dependence, in which the AI's future trajectory is constrained by its earlier termination decisions [7].

Second, abandonment is the primary generator of regret within materials discovery campaigns. The framework identifies regret not as an incidental emotional residue but as a quantifiable epistemic cost that accumulates whenever a later independent study or serendipitous re-exploration reveals value in a previously terminated path. Regret tracking, as proposed in the detection principles, therefore serves as the natural bridge between abandonment and this broader regret failure mode [3, 28].

Third, path abandonment represents the moment when the exploration–exploitation trade-off tilts decisively—and often prematurely—toward exploitation. Active-learning systems are designed to navigate this trade-off, yet the abandonment mechanisms (particularly exploration decay) cause the balance to collapse into pure exploitation of the current local basin. The framework, therefore, positions abandonment as the observable symptom of an exploration–exploitation imbalance rather than an independent pathology [21, 22].

Fourth, abandonment contributes directly to epistemic debt. By terminating paths before sufficient evidence has been gathered, the system accrues unexamined possibilities that remain hidden from the collective scientific record. This debt compounds because downstream models are trained on biased datasets that exclude the abandoned regions, thereby propagating the initial error. The relation to epistemic debt underscores why mitigation must include path revisitation and regret-aware stopping: only by periodically repaying the debt can the community avoid systematic blind spots in materials knowledge [5, 7].

These interrelations demonstrate that addressing scientific path abandonment is not a narrow technical fix but a leverage point for improving the robustness of the entire AI-guided discovery ecosystem.

Implications for Materials AI Practice

The identification of scientific path abandonment as a distinct failure mode carries concrete implications for how materials researchers, reviewers, and the broader community should adapt their practices.

For authors, the framework proposes three mandatory reporting standards: (a) explicit documentation of every abandonment decision within the methods section, including the governing stopping rule and the mechanism inferred by the abandonment audit; (b) justification of the chosen stopping criteria against the evidentiary threshold defined in Definition 1; and (c) inclusion of a regret ledger summarizing the estimated opportunity cost of terminated paths. These requirements transform abandonment from an opaque background process into a transparent scientific choice [3, 7].

For reviewers, the framework recommends two targeted interrogation prompts: (a) “Does the manuscript demonstrate that abandonment decisions were subjected to counterfactual analysis or diversity monitoring?” and (b) “Are the stopping criteria justified against the risk of false-negative cascades or local-optima traps?” Such questions elevate abandonment from an assumed optimization detail to a central validity concern.

For the community at large, three systemic changes are articulated: (a) development of standardized abandonment ontologies and metadata schemas that can be shared across laboratories and repositories; (b) creation of open-source path-tracking tools that visualize abandonment surfaces in real time; and (c) establishment of longitudinal studies that mine published campaigns for patterns of premature termination. These community-level interventions would convert the present conceptual framework into an operational infrastructure that makes scientific path abandonment both detectable and mitigable at scale [4, 21].

Taken together, these practice-level implications shift materials AI from a paradigm that optimizes for speed toward one that optimizes for epistemic responsibility.

Conclusion

Scientific path abandonment has been articulated in this analysis as a systemic failure mode in AI-guided materials research: the premature termination of promising research directions before sufficient evidence has been gathered to determine their true promise. By distinguishing abandonment from rational stopping, pruning, and switching; by dissecting its four core mechanisms of impatience, overconfidence, narrow optimization, and exploration decay; by cataloging its four primary types; and by delineating the resulting failure modes of local optima traps, false-negative cascades, exploration starvation, and regret amplification, the framework offers a comprehensive diagnostic and remedial vocabulary. Detection principles centered on audits, counterfactuals, diversity monitoring, and regret tracking, together with mitigation principles of extended exploration, exploration reserves, delayed abandonment, path revisitation, and regret-aware stopping, provide actionable levers for prevention. The relationships to path dependence, regret, the exploration–exploitation imbalance, and epistemic debt further embed abandonment within the broader landscape of AI failure modes.

The ultimate call is for abandonment-aware materials AI. This approach resists premature path termination not by eliminating stopping rules, but by ensuring that every termination is based on sufficient evidence rather than algorithmic convenience. Only through such vigilance can the promise of autonomous discovery be realized without sacrificing the very scientific serendipity and long-horizon exploration that materials innovation demands. Future work should focus on embedding these principles into widely adopted open-source platforms, thereby transforming path abandonment from an invisible risk into a visible and manageable feature of responsible AI-guided science.

Acknowledgements

None

Conflict of interest

None

Ethics statement

None

Financial support

None

Received: 12 Apr 2025 Revised: 30 Jun 2025 Accepted: 05 Aug 2025

Rights and permissions

References

Antoniuk ER, Li P, Keilbart N, Weitzner S, Kailkhura B, Hiszpanski AM. Active learning enables generation of molecules that advance the known Pareto front. npj Comput Mater. 2026;12(1):81.

Ko TW, Deng B, Nassar M, Barroso-Luque L, Liu R, Qi J, et al. Materials graph library (MatGL), an open-source graph deep learning library for materials science and chemistry. npj Comput Mater. 2025;11(1):253.

Maqsood A, Chen C, Jacobsson TJ. The future of material scientists in an age of artificial intelligence. Adv Sci. 2024;11(19):2401401.

Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature. 2018;559(7715):547-55.

Schmidt J, Marques MR, Botti S, Marques MA. Recent advances and applications of machine learning in solid-state materials science. npj Comput Mater. 2019;5(1):83.

Zunger A. Inverse design in search of materials with target functionalities. Nat Rev Chem. 2018;2(4):0121.

Montoya JH, Aykol M, Anapolsky A, Gopal CB, Herring PK, Hummelshøj JS, et al. Toward autonomous materials research: Recent progress and future challenges. Appl Phys Rev. 2022;9(1).

Maxson T, Soyemi A, Zhang X, Chen BW, Szilvási T. MS25: Materials science-focused benchmark data set for machine learning interatomic potentials. J Chem Inf Model. 2025;65(15):8097-112.

Schmidt J, Cerqueira TF, Romero AH, Loew A, Jäger F, Wang HC, et al. Improving machine-learning models in materials science through large datasets. Mater Today Phys. 2024;48:101560.

Shoghi R, Morand L, Helm D, Hartmaier A. Optimizing machine learning yield functions using query-by-committee for support vector classification with a dynamic stopping criterion. Comput Mech. 2024;74(2):447-66.

Zou Y. Smart material optimization using reinforcement learning in multi-dimensional self-assembly. Front Mater. 2025;12:1526892.

Tian Y, Li T, Pang J, Zhou Y, Xue D, Ding X, et al. Materials design with target-oriented Bayesian optimization. npj Comput Mater. 2025;11(1):209.

Rodemann J, Augustin T. Imprecise Bayesian optimization. Knowl Based Syst. 2024;300:112186.

Schoepfer AA, Weinreich J, Laplaza R, Waser J, Corminboeuf C. Cost-informed Bayesian reaction optimization. Digit Discov. 2024;3(11):2289-97.

Al-Rimy BA, Saeed F, Al-Sarem M, Albarrak AM, Qasem SN. An adaptive early stopping technique for DenseNet169-based knee osteoarthritis detection model. Diagnostics. 2023;13(11):1903.

Hernandez JG, Saini AK, Ghosh A, Moore JH. The tree-based pipeline optimization tool: Tackling biomedical research problems with genetic programming and automated machine learning. Patterns. 2025;6(7).

Zanjani Foumani Z, Bostanabad R. Cost-aware Bayesian optimization with automatic stop condition under multi-fidelity constraints and data. J Mech Des. 2026;148(2):021711. https://doi.org/10.1115/1.4070314.

Ishibashi H, Karasuyama M, Takeuchi I, Hino H. A stopping criterion for Bayesian optimization by the gap of expected minimum simple regrets. In: International Conference on Artificial Intelligence and Statistics. PMLR; 2023. p. 6463-97.

König L, Zitzmann S, Fütterer T, Campos DG, Scherer R, Hecht M. An evaluation of the performance of stopping rules in AI-aided screening for psychological meta-analytical research. Res Synth Methods. 2024;15(6):1120-46.

Callaghan M, Müller-Hansen F, Bond M, Hamel C, Devane D, Kusa W, et al. Computer-assisted screening in systematic evidence synthesis requires robust and well-evaluated stopping criteria. Syst Rev. 2024;13(1):284.

Nair AS, Foppa L. A critical examination of active learning workflows in materials science. arXiv preprint arXiv:2601.05946. 2026 Jan 9.

Aqib M, Ravikiran K, Li L, Prasad V. A multi-fidelity active learning framework for accelerated alloy design. Mater Des. 2026:115520.

Goodall RE, Parackal AS, Faber FA, Armiento R, Lee AA. Rapid discovery of stable materials by coordinate-free coarse graining. Sci Adv. 2022;8(30):eabn4117.

Vasanthakumari P, Zhu Y, Brettin T, Partin A, Shukla M, Xia F, et al. A comprehensive investigation of active learning strategies for conducting anti-cancer drug screening. Cancers. 2024;16(3):530.

Nacar EN, Erdebilli B, Eraslan E. Toward green manufacturing: A heuristic hybrid machine learning framework with PSO for scrap reduction. Sustainability. 2025;17(20):9106.

Schweitzer F, Habel L, Canaviri Vilca OJ, Kulzer T, Wenzel S. Advancing production planning in the automotive industry: Machine learning-enhanced simulation optimization using metamodels. J Simul. 2025:1-7.

Witkowski T. Implicit switching-cost regularization in supervised trading signal classification. https://doi.org/10.21203/rs.3.rs-8947769/v1.

Xie Q, Cai L, Terenin A, Frazier PI, Scully Z. Cost-aware stopping for Bayesian optimization. arXiv preprint arXiv:2507.12453. 2025 Jul 16.

Yoosefzadeh-Najafabadi M, Tulpan D, Eskandari M. Application of machine learning and genetic optimization algorithms for modeling and optimizing soybean yield using its component traits. PLoS One. 2021;16(4):e0250665.