ORIGINAL RESEARCH **Open access**

Conceptual Foundations for Scientific Contestability in AI-Driven Materials Decisions

Mohammed Al-Farsi[1*], Salim Al-Harthy[1], Nasser Al-Rawahi[2]

Abstract

In contemporary materials science, artificial intelligence systems increasingly generate high-stakes decisions—recommending specific compositions for synthesis, prioritizing experimental campaigns, or endorsing candidate structures for further validation—yet these systems typically provide no structured pathway for scientists to challenge, appeal, or revise the outputs when they appear erroneous or misaligned with domain knowledge. Scientific contestability is defined here as the capacity for scientists to formally challenge, appeal, or request revision of AI-generated decisions through transparent procedures that guarantee meaningful reconsideration grounded in epistemic and procedural norms. This principle matters profoundly in materials AI because erroneous recommendations can waste substantial laboratory resources, delay critical technological advances, exacerbate epistemic uncertainty inherent to data-driven predictions, undermine scientific pluralism by privileging singular algorithmic perspectives, and violate basic requirements of procedural justice for researchers whose careers and discoveries depend on these outputs. The present framework articulates five interlocking components—contestation triggers, mechanisms, review processes, decision revision pathways, and record keeping—that together transform contestability from an abstract ideal into a practical design requirement for materials AI platforms. By embedding contestability at the core of system architecture, the framework offers concrete implications for designers, researchers, and institutions, ensuring that AI-assisted materials discovery remains epistemically robust, democratically accountable, and aligned with the self-correcting ethos of science.

Keywords Scientific contestability, Algorithmic contestability, AI-driven materials discovery, Procedural justice in AI, Epistemic uncertainty in machine learning, Contestable materials AI

*Correspondence:
Mohammed Al-Farsi
mohammed.alfarsi@gmail.com
[1] Department of Intelligent Materials Analytics, Sultan Qaboos University, Muscat, Oman
[2] Department of AI-Based Materials Systems, German University of Technology in Oman, Muscat, Oman

Introduction

Materials AI systems now routinely make consequential decisions that shape the trajectory of scientific research: suggesting which novel perovskite composition should be synthesized next, which synthesis route deserves experimental prioritization, or which computationally predicted crystal structure merits immediate laboratory validation. Yet when these decisions prove flawed—perhaps because the underlying training data overlooked a critical failure mode, or because the model's latent assumptions clash with newly acquired experimental evidence—scientists currently lack any formalized mechanism to contest, appeal, or demand revision of the AI output. The absence of contestability mechanisms leaves researchers in a position of passive acceptance, even when domain expertise indicates that the recommendation is suboptimal or outright incorrect. This paper introduces scientific contestability as a foundational design principle for AI

systems deployed in materials science, arguing that the ability to challenge AI-generated decisions is not an optional add-on but an essential epistemic safeguard [1-3].

Butler *et al.* [4] have demonstrated that machine learning has become indispensable for navigating the vast combinatorial space of possible materials. Yet, the very power of these models creates new vulnerabilities when their outputs go unchallenged. Hendry, writing on contestability in algorithmic systems more broadly, emphasizes that without structured pathways for disagreement, automated decision-making risks becoming unaccountable even in domains where error carries high scientific and economic costs [1]. The present work builds directly on these insights by adapting contestability concepts to the unique demands of materials discovery, where decisions are not merely administrative but shape the very frontier of knowledge production. Binns connects algorithmic accountability to public reason [2], a linkage that becomes especially salient when AI systems influence which research questions receive resources and which are sidelined.

Current materials AI platforms, while impressive in predictive accuracy, remain silent on the question of contestation. A scientist who suspects that an AI-recommended thermoelectric candidate will fail under realistic processing conditions has no formal channel through which to register that disagreement, supply counter-evidence, or trigger a reconsideration process. This gap is not merely technical; it is epistemic and normative. The framework developed here, therefore, begins by defining scientific contestability, demonstrates why it is indispensable in materials contexts, proposes a five-component conceptual architecture, and elaborates each component in detail. In doing so, the paper establishes contestability as a necessary complement to existing emphases on accuracy, efficiency, and scalability in AI-driven materials research.

Defining Scientific Contestability

Scientific contestability can be formally stated as follows:

Definition 1: Scientific contestability is the capacity for scientists to challenge, appeal, or request revision of AI-generated decisions within materials discovery pipelines, supported by transparent procedures that provide meaningful opportunity for reconsideration grounded in epistemic evidence, methodological scrutiny, or normative disagreement.

This definition deliberately foregrounds agency: contestability is not passive transparency but an active, procedurally protected right to push back against algorithmic outputs. It must be distinguished from several neighboring concepts that, while valuable, are insufficient on their own. Explainability, as discussed extensively in the literature on interpretable machine learning, concerns the extent to which a model's internal logic can be made intelligible to human users. Yet knowing why a model predicted a particular bandgap value does not automatically equip a scientist to contest that prediction when new physical intuition suggests the value is implausible. Interpretability similarly focuses on making model internals transparent—through attention maps, feature importance scores, or surrogate models—but stops short of institutionalizing a pathway for revision once the output has been issued.

Accountability, by contrast, typically addresses the attribution of responsibility after harm has occurred, as Binns articulates in linking algorithmic systems to public reason [2]. Contestability, however, operates upstream: it enables proactive correction rather than retrospective blame. Feedback mechanisms, common in interactive AI interfaces, allow users to provide input that may refine future models. Yet, they lack the formal, time-bound, and outcome-oriented structure that contestability demands. A casual "this prediction seems off" comment differs fundamentally from a documented appeal that triggers mandatory review and possible override [5-10].

Table 1 analytically differentiates scientific contestability from adjacent concepts, demonstrating its unique role as a procedural and intervention-oriented extension of existing AI governance principles.

Table 1. Analytical differentiation of scientific contestability from adjacent AI governance concepts

Concept	Core function	Temporal orientation	Level of intervention	Limitation without contestability	Contribution of contestability
Explainability	Makes model reasoning interpretable	Pre-decision or post-hoc	Cognitive/epistemic	Does not enable challenge or revision	Converts understanding into an actionable challenge
Interpretability	Reveals internal model structure	Pre-decision	Technical/model-level	Stops at transparency	Enables procedural intervention beyond model inspection
Accountability	Assigns responsibility for outcomes	Post-decision	Institutional/legal	Reactive and retrospective	Introduces proactive correction mechanisms
Feedback mechanisms	Collects user input for improvement	Iterative/long-term	Informal system interaction	Lacks procedural force and guarantees	Formalizes structured, outcome-oriented appeals
Scientific contestability	Enables formal challenge and revision	Real-time / pre-implementation	Procedural and epistemic	—	Integrates explanation, accountability, and procedural justice into actionable system design

Zarsky and others have noted that the right to contest decisions is especially critical in high-stakes domains [11]; the same logic applies with equal force to scientific decisions whose downstream consequences include misallocated grant funding, abandoned research lines, or premature publication of flawed predictions. Scientific contestability, therefore, integrates elements of procedural fairness, epistemic humility, and institutional design into a single, actionable principle tailored to AI-assisted materials science.

Why Contestability Matters in Materials AI

Contestability is not a luxury but a necessity for four interlocking reasons that reflect the distinctive character of materials discovery.

Argument 1: High stakes

Materials recommendations frequently commit substantial laboratory resources—equipment time, reagents, and personnel hours—to synthesis and characterization campaigns that may ultimately fail. As Montoya *et al.* [7] have shown in their analysis of autonomous research platforms, the financial and temporal costs of pursuing AI-endorsed candidates are considerable. Without contestability, scientists are compelled to accept these costs even when domain expertise flags potential failure modes.

Argument 2: Epistemic uncertainty

Machine learning models for materials properties operate under inherent uncertainty arising from incomplete training data, extrapolation beyond known chemical spaces, and the stochastic nature of many physical phenomena. Butler *et al.* highlight that even state-of-the-art models for molecular and materials science retain significant predictive error bands [4]. Contestation provides a systematic mechanism for injecting new evidence or alternative interpretations precisely when uncertainty is highest.

Argument 3: Scientific pluralism

Different research groups legitimately prioritize different criteria—thermodynamic stability versus kinetic accessibility, cost versus performance, or sustainability versus raw functionality. Schmidt and co-authors document the rapid expansion of machine learning applications across solid-state materials [5], yet these models inevitably encode singular optimization objectives. Contestability preserves space for pluralistic scientific judgment rather than allowing algorithmic monoculture to dominate.

Argument 4: Procedural justice

Researchers whose careers depend on the acceptance or rejection of AI recommendations deserve due process. Drawing on broader discussions of automated decision-making, contestability ensures that affected scientists are not merely passive recipients but active participants in the epistemic process, thereby upholding norms of fairness and legitimacy in data-driven science.

A Framework for Scientific Contestability

This paper proposes a conceptual framework for scientific contestability in AI-driven materials decisions organized around five interdependent components. The framework can be conceptualized as a cyclical process flow: an AI-generated decision first encounters a contestation trigger; if activated, the trigger invokes a designated mechanism; the mechanism routes the challenge into a structured review process; the review culminates in a decision revision outcome; and finally, all steps feed into record keeping that refines both the AI model and future contestation procedures. This flow ensures that contestability is not an afterthought but an embedded architectural feature.

Figure 1 presents the hierarchical architecture of scientific contestability, illustrating how AI-generated decisions are systematically subjected to structured triggers, mechanisms, review processes, and revision pathways.

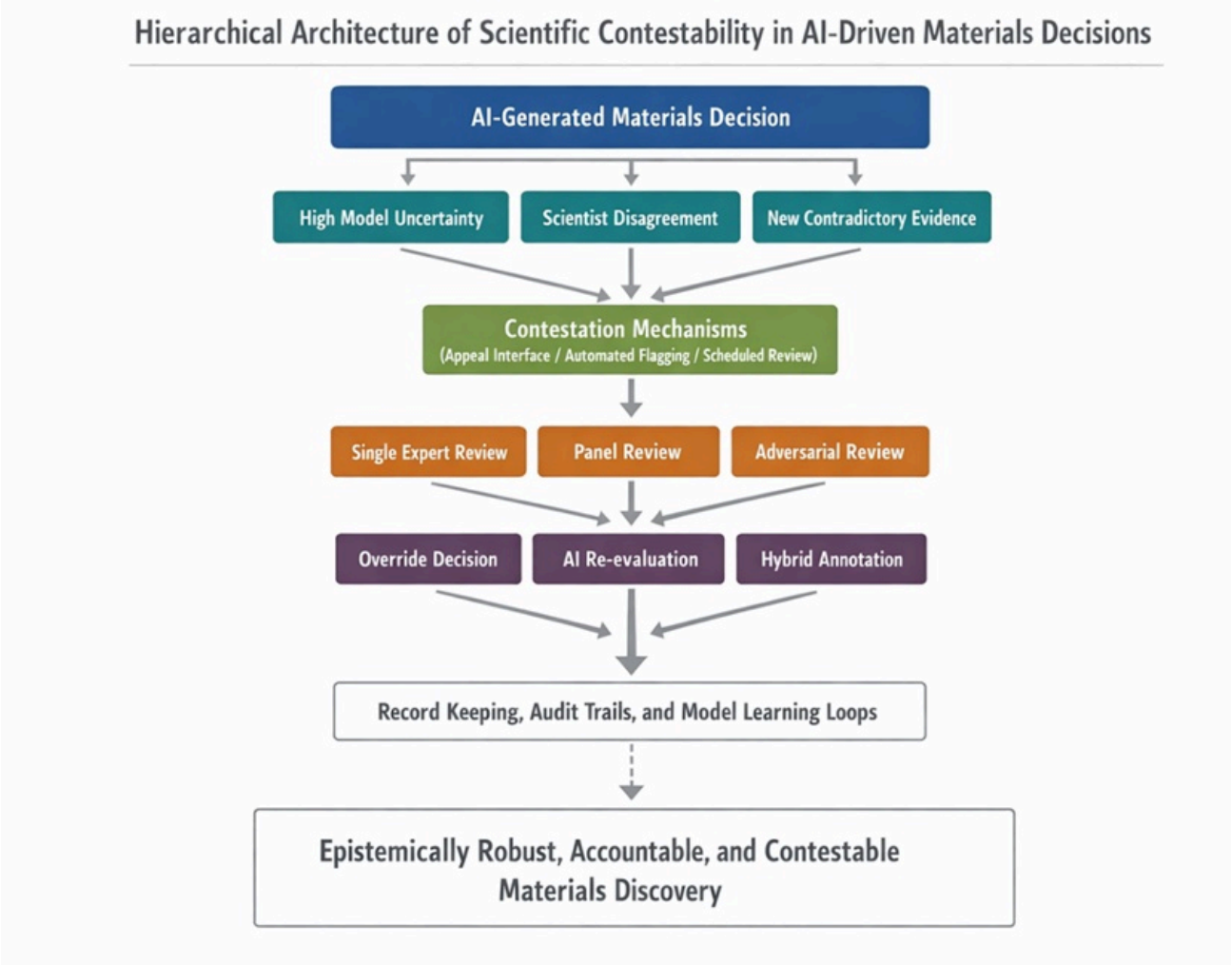

Figure 1. The hierarchical architecture of scientific contestability

The five components are deliberately modular so that they can be implemented across different platforms—whether cloud-based autonomous laboratories or desktop materials design tools—while remaining conceptually unified. Each component draws on insights from algorithmic contestability literature while adapting them to the epistemic demands of materials science. For instance, Hendry's emphasis on contestability by design finds concrete realization here through explicit triggers and mechanisms [1], while Binns' focus on public reason informs the review and revision stages [2].

Components of the Framework

The framework's first component is the contestation trigger. Three primary conditions activate contestation: (a) the AI's internal uncertainty metric exceeds a pre-defined, domain-calibrated threshold; (b) the scientist explicitly registers substantive disagreement with the output based on prior experimental knowledge or alternative theoretical models; and (c) new contradictory evidence—such as a recently published experimental result or an independent computational validation—becomes available during the decision window. These triggers ensure that contestation is neither frivolous nor automatic but epistemically grounded.

The second component comprises the contestation mechanism. Implementation options include (a) a dedicated interface element (for example, an "Appeal Decision" button integrated into the AI dashboard) that launches a structured appeal form; (b) automated flagging routines that detect statistically anomalous predictions and proactively invite contestation;

and (c) periodic review cycles built into collaborative platforms that surface contested or high-uncertainty recommendations for collective scrutiny. These mechanisms lower the practical barrier to contestation while preserving scientific rigor.

The third component is the review process. Viable configurations encompass (a) review by a single qualified human expert drawn from a pre-approved pool of domain specialists; (b) panel review involving multiple researchers to capture pluralistic perspectives; and (c) an adversarial process in which the original AI output is defended by one agent while the contesting scientist presents counter-arguments. The choice of process can scale with the decision's stakes, ensuring proportionality.

The fourth component addresses decision revision. Possible outcomes are (a) full override by human judgment when contestation is upheld; (b) a request for the AI system to reconsider its recommendation after incorporating the contesting evidence; and (c) hybrid resolution in which the AI output is retained but annotated with the contestation record and confidence adjustment. Revision pathways, therefore, range from conservative to transformative.

The fifth component is record keeping. Essential functions include (a) systematic tracking of every contestation outcome together with the evidence supplied; (b) learning loops that feed contested cases back into model retraining or fine-tuning; and (c) audit trails that allow institutions to monitor contestation patterns, identify systemic biases, and refine contestability procedures over time. Robust record keeping transforms individual challenges into collective epistemic improvement.

Types of Contestation

Scientific contestability in AI-driven materials decisions operates through four distinct yet complementary types of contestation, each targeting a different dimension of potential misalignment between algorithmic output and scientific judgment. These types provide a typology that allows researchers to articulate precisely why an AI recommendation warrants reconsideration, thereby channeling epistemic disagreement into structured, productive pathways rather than diffuse frustration. By distinguishing these categories, the framework ensures that contestation remains targeted, evidence-based, and proportionate to the nature of the challenge, drawing on the procedural mechanisms outlined in the five-component architecture.

Type 1: Factual contestation is defined as the formal challenge to the empirical or predictive correctness of an AI-generated output, where the scientist asserts that the stated material property, stability, or performance metric is factually inaccurate given available evidence. In materials science, a concrete example arises when an AI platform recommends a particular metal-organic framework composition as possessing an exceptionally high CO_2 uptake capacity at room temperature, yet the contesting researcher possesses unpublished or recently acquired isotherm data indicating that the framework collapses under humid conditions, rendering the prediction factually untenable. Cobbe and Veale emphasize that factual contestation must be supported by verifiable counter-evidence [8], and the appropriate review process for this type is typically a panel review incorporating independent computational validation alongside the new experimental data, ensuring that the factual claim is adjudicated through rigorous cross-verification rather than subjective opinion.

Type 2: Methodological contestation targets the underlying modeling assumptions, training protocols, or algorithmic choices that produced the recommendation. Here, the scientist does not necessarily dispute the numerical output but argues that the method itself is ill-suited to the problem domain—for instance, when a graph neural network trained predominantly on oxide perovskites is applied to predict phase stability in halide perovskites without adequate transfer learning, leading to systematic extrapolation errors. Montoya *et al.* [7] highlight the methodological pitfalls inherent in autonomous materials platforms, underscoring why such contestation is essential. The suitable review process for methodological contestation is an adversarial format in which one reviewer defends the original model architecture. At the same time, the contesting scientist presents alternative methodological benchmarks, such as density-functional-theory

cross-checks or alternative featurization schemes, thereby exposing whether the AI's methodological scaffolding aligns with domain standards [12-19].

Type 3: Interpretive contestation arises when scientists agree on the raw prediction yet disagree on its scientific meaning or implications for downstream research. A typical materials example occurs when an AI identifies a new cathode material with a theoretically high voltage window; one group interprets this as immediately actionable for battery prototyping, while another interprets the same output as requiring urgent caution because the voltage window exceeds the electrochemical stability limit of common electrolytes, altering the practical interpretation dramatically. Wachter and co-authors note that interpretive layers are often overlooked in AI systems [13], yet they carry substantial consequences for research prioritization. The review process best suited to interpretive contestation is single-expert review by a senior domain specialist who can weigh competing interpretations against a broader literature context, providing a calibrated judgment that respects interpretive pluralism without descending into relativism.

Type 4: Normative contestation challenges the values, priorities, or optimization objectives implicitly embedded within the AI decision. For example, an AI system optimized exclusively for maximum piezoelectric coefficient may recommend a lead-containing material. Yet, the contesting researcher argues that sustainability and toxicity constraints—normative considerations increasingly central to green materials design—should override raw performance metrics. Butler *et al.* [4] implicitly acknowledge such value trade-offs when discussing the broader societal embedding of machine-learning tools in materials science. The appropriate review process here is a hybrid panel that includes both technical experts and institutional ethicists or sustainability officers, ensuring that normative contestation is adjudicated through explicit deliberation on value weights rather than hidden in black-box optimization functions.

Collectively, these four types illustrate how contestability transforms vague unease into precise, actionable epistemic interventions, each mapped to a tailored review pathway within the overarching framework.

Table 2 consolidates the four types of contestation by mapping each to its evidentiary requirements, review structure, and associated epistemic risks, thereby clarifying their distinct operational roles.

Table 2. Structured mapping of contestation types to evidence requirements, review designs, and epistemic risks

Contestation type	Primary target	Required evidence	Appropriate review design	Key epistemic risk if absent	Outcome sensitivity
Factual contestation	Empirical accuracy of prediction	Experimental data and validation studies	Panel review with cross-verification	Propagation of false predictions into experiments	High
Methodological contestation	Model assumptions and training design	Benchmark comparisons and alternative models	Adversarial review	Systematic model bias and invalid extrapolation	Very High
Interpretive contestation	Meaning and implications of outputs	Literature context and domain reasoning	Single expert review	Misguided research prioritization	Moderate
Normative contestation	Embedded values and optimization criteria	Policy constraints and ethical considerations	Hybrid panel (technical + ethical)	Reinforcement of harmful or narrow optimization goals	High

Relation to Existing Concepts

Scientific contestability does not emerge as an isolated normative addition to AI governance; rather, it is best understood as a synthetic extension that consolidates and operationalizes several foundational concepts in AI ethics and interpretability. Its contribution lies precisely in transforming these largely abstract or loosely connected principles into a cohesive, action-oriented framework tailored to scientific discovery contexts.

Most prominently, contestability builds upon—yet decisively moves beyond—explainability. Explainability techniques, including feature attribution methods, saliency mapping, and counterfactual generation, are designed to render model behavior intelligible to human users. These tools illuminate the internal logic of AI systems, thereby reducing epistemic opacity. However, they remain fundamentally descriptive rather than interventionist. A researcher may fully understand why a model recommends a specific alloy composition or processing pathway, yet still lack any formal mechanism to dispute or override that recommendation when it conflicts with empirical intuition or domain expertise. In this sense, explainability without contestability risks becoming epistemically inert. As Wachter suggests, counterfactual explanations can indeed support contestation by clarifying decision boundaries [18], but they do not themselves constitute contestation. Contestability, by contrast, introduces a procedural dimension: it establishes the right—and crucially, the structured means—to challenge, revise, and potentially overturn AI outputs within an institutional framework [20-24].

Contestability is equally intertwined with the concept of accountability, yet it reorients its temporal and functional scope. Binns frames algorithmic accountability in terms of public reason, emphasizing that affected parties must be able to question and justify automated decisions [2]. While this perspective is essential, it is often operationalized retrospectively, focusing on assigning responsibility after harm or error has occurred. Scientific contestability extends this logic upstream. It embeds mechanisms for critique and revision directly into the decision-making pipeline, enabling intervention before flawed recommendations translate into wasted resources, failed experiments, or misleading conclusions. In this way, contestability transforms accountability from a reactive practice into a proactive safeguard, ensuring that the epistemic integrity of AI-assisted research is maintained in real time rather than reconstructed after the fact.

The framework further aligns with—and concretizes—the principles of procedural justice in automated systems. Procedural justice emphasizes fairness not only in outcomes but in the processes that generate those outcomes. Within AI-assisted materials discovery, contestability ensures that researchers are not reduced to passive recipients of algorithmic authority but remain active epistemic agents with the right to a fair hearing. Brkan's analysis of the right to challenge automated decisions under GDPR highlights the necessity of procedural safeguards such as transparency, reviewability, and the opportunity for human intervention [9]. Scientific contestability adapts and extends these safeguards into the domain of research practice, where the stakes are not only economic or legal but epistemic—shaping what is accepted as valid knowledge. By embedding contestation rights within scientific workflows, the framework preserves the dialogical and adversarial character of scientific inquiry.

Finally, contestability functions as a formalized mechanism of systematic error correction, elevating what are often informal feedback practices into structured, auditable processes. Traditional scientific progress relies on iterative critique, replication, and peer review; however, AI systems can inadvertently bypass these mechanisms by presenting outputs with an aura of computational authority. Hendry underscores that contestability is indispensable precisely because algorithmic systems are inherently fallible [1]. This fallibility is amplified in materials science, where high-dimensional chemical and structural spaces introduce profound uncertainty and sparse empirical validation. By institutionalizing contestation pathways, the framework ensures that errors are not only detectable but correctable within a transparent and accountable system.

Taken together, scientific contestability integrates explainability, accountability, procedural justice, and error correction into a unified design principle. Its novelty lies not in introducing entirely new ethical commitments, but in orchestrating existing ones into a coherent operational architecture that is greater than the sum of its parts—one capable of sustaining trustworthy, adaptive, and epistemically robust AI-driven discovery.

Implications for Materials AI Practice

The adoption of scientific contestability is not merely a conceptual refinement; it entails concrete and far-reaching changes in how materials AI systems are designed, used, and governed. Its implications unfold across three primary stakeholder groups: system designers, individual researchers, and institutional actors [25-29].

For system designers, contestability necessitates a paradigm shift from performance-centric optimization toward governance-aware architecture. Rather than treating contestation as an afterthought or external add-on, designers must embed it as a core system functionality from the earliest stages of development. This includes the integration of dedicated contestation interfaces that allow users to formally challenge model outputs, as well as uncertainty-sensitive mechanisms that proactively flag recommendations requiring human review. In addition, comprehensive logging and traceability infrastructures must be implemented to record not only model decisions but also the history of contestations, revisions, and outcomes. Such records enable both internal auditing and external scrutiny, fostering transparency and continuous improvement. As Selbst and Barocas argue, incorporating contestability by design mitigates downstream harms and aligns system behavior with broader social and ethical expectations [15]. In the context of materials AI, this translates into iterative co-evolution between models and users, where contestation data becomes a critical input for retraining and interface refinement.

For individual researchers, scientific contestability redefines the epistemic role of the human expert. Instead of passively accepting algorithmic outputs as authoritative, researchers are expected to engage critically and constructively with AI-generated recommendations. This involves actively invoking contestation mechanisms when discrepancies arise between model outputs and domain knowledge, as well as providing substantiated evidence—such as prior literature, experimental data, or theoretical reasoning—to support their challenges. Moreover, researchers may be called upon to participate in peer-like review processes, evaluating contestations submitted by others and contributing to collective decision-making. This shift represents a cultural transformation, reinforcing the inherently self-correcting nature of scientific practice. By encouraging active engagement, contestability helps prevent the silent accumulation of errors and biases that can undermine the promise of autonomous research systems, as highlighted by Montoya *et al.* [7].

For research institutions, funding agencies, and journal editors, the implications are both organizational and infrastructural. Institutions must develop standardized protocols for contestation, ensuring consistency and fairness across different projects and platforms. This includes establishing standing review panels composed of interdisciplinary experts capable of adjudicating contested AI outputs. Additionally, dedicated resources must be allocated to support the technical and administrative infrastructure required for contestable systems, including data storage, audit tools, and user training programs. Importantly, contestation records should be recognized as valuable epistemic artifacts rather than mere administrative byproducts. Aggregated contestation data can reveal systematic patterns of model failure, bias, or uncertainty, thereby informing targeted model improvements and guiding future research priorities.

At a broader level, embedding contestability within institutional practices helps safeguard the pluralistic and evidence-driven character of materials science. It ensures that AI systems augment rather than displace human judgment, preserving the diversity of perspectives that is essential for innovation. By formalizing the right to challenge and revise algorithmic outputs, institutions can create an environment in which AI serves as a collaborator in discovery rather than an unchallengeable authority.

Conclusion

This paper has introduced scientific contestability as a foundational design principle for AI-driven materials decisions, formally defined it, demonstrated its necessity across high-stakes, epistemically uncertain, and pluralistic research contexts, proposed a five-component conceptual framework, elaborated each component with concrete sub-elements, and delineated four distinct types of contestation together with their interrelations to explainability, accountability, and procedural justice. The framework—centered on contestation triggers, mechanisms, review processes, decision revision

pathways, and record keeping—offers a modular yet coherent architecture capable of implementation across diverse materials AI platforms.

Scientific contestability must therefore become a standard, non-negotiable feature of all materials AI systems, ensuring that the remarkable predictive power of machine learning remains tethered to the critical, corrective, and pluralistic spirit of scientific inquiry. Only through such contestability can AI-assisted materials discovery fulfill its promise without sacrificing the epistemic humility and procedural fairness that define rigorous science.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 04 Jan 2025 Revised: 13 Feb 2025 Accepted: 21 Apr 2025
Published online: 18 July 2025

References

Leofante F, Ayoobi H, Dejl A, Freedman G, Gorur D, Jiang J, et al. Contestable AI needs computational argumentation. arXiv preprint arXiv:2405.10729. 2024 May 17.

Binns R. Algorithmic accountability and public reason. Philos Technol. 2018;31(4):543-56.

Yonathan M. Mathematical foundations for machine learning: A programming-enhanced educational framework for undergraduate education. Available from: https://www.ssrn.com/abstract=5367023

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).

Vaccaro K, Karahalios K, Mulligan DK, Kluttz D, Hirsch T. Contestability in algorithmic systems. In: Companion Publication of the 2019 Conference on Computer Supported Cooperative Work and Social Computing; 2019. p. 523-7.

Tretter M. Opportunities and challenges of AI-systems in political decision-making contexts. Front Polit Sci. 2025;7:1504520.

Burgun A, Do V, Elamrani A, Kirat T, Lang J, Poibeau T. Challenges in AI-assisted decision-making: Insights from the medical and the legal domains. 2025.

Board DI. AI principles: Recommendations on the ethical use of artificial intelligence by the department of defense: Supporting document. United States Department of Defense. 2019.

Schnitzer M, Montag F. Preserving contest ability in the data economy. Compet Law Policy Debate. 2019;5(2):55-65.

McInerney T. The algorithmic construction of epistemic injustice. In: Socio-legal studies on epistemic injustice and spaces and places. Cham: Springer Nature Switzerland; 2026. p. 123-54.

Suksi M. Administrative due process when using automated decision-making in public administration: Some notes from a Finnish perspective. Artif Intell Law. 2021;29(1):87-110.

Orphanou K, Otterbacher J, Kleanthous S, Batsuren K, Giunchiglia F, Bogina V, et al. Mitigating bias in algorithmic systems—a fish-eye view. ACM Comput Surv. 2022;55(5):1-37.

Cruz-Aguilar MA. The epistemic revolution of AI: Reconfiguring the foundations of scientific knowledge. AI Soc. 2025:1-7.

Chen SS. The dawn of AI generated contents: Revisiting compulsory mediation and IP disputes resolution. Contemp Asia Arb J. 2023;16:301.

Moreira C, Palatkina A, Braca D, Walsh DM, Leihn PJ, Chen F, et al. Explainable AI systems must be contestable: Here's how to make it happen. arXiv preprint arXiv:2506.01662. 2025 Jun 2.

Dignum V, Michael L, Nieves JC, Slavkovik M, Suarez J, Theodorou A. Contesting black-box AI decisions. In: Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS); 2025. p. 2854-8.

Lyons H, Velloso E, Miller T. Conceptualising contestability: Perspectives on contesting algorithmic decisions. Proc ACM Hum Comput Interact. 2021;5(CSCW1):1-25.

Mukherjee A, Chang HH. Agentic AI: Autonomy, accountability, and the algorithmic society. arXiv preprint arXiv:2502.00289. 2025 Feb 1.

Liedong TA. The liability of tribe in corporate political activity: Ethical implications for political contestability. J Bus Ethics. 2022;181(3):623-44.

Reeves-McLaren N, Christensen SM. Data integrity in materials science in the era of AI: Balancing accelerated discovery with responsible science and innovation. J Mater Chem A. 2026;14(1):276-83.

Cath C, Wachter S, Mittelstadt B, Taddeo M, Floridi L. Artificial intelligence and the good society: The US, EU, and UK approach. Sci Eng Ethics. 2018;24(2):505-28.

Pohlhaus Jr G. Knowing without borders and the work of epistemic gathering. In: Decolonizing feminism: Transnational feminism and globalization; 2017. p. 37-54.

Landau S, Dempsey JX, Kamar E, Bellovin SM, Pool R. Challenging the machine: Contestability in government AI systems. arXiv preprint arXiv:2406.10430. 2024 Jun 14.

Alfrink K, Keller I, Kortuem G, Doorn N. Contestable AI by design: Towards a framework. Minds Mach. 2023;33(4):613-39.

Singh AV, Rosenkranz D, Ansari MH, Singh R, Kanase A, Singh SP, et al. Artificial intelligence and machine learning empower advanced biomedical material design to toxicity prediction. Adv Intell Syst. 2020;2(12):2000084.

Delacroix S. Diachronic interpretability and machine learning systems. J Cross Discip Res Comput Law. 2022;1(2).