ORIGINAL RESEARCH **Open access**

The Problem of Representational Harm in Materials Dataset Construction

George Brown[1], Michael Taylor[1*], Sarah Wilson[2], Olivia Harris[1]

Abstract

Representational harm in materials dataset construction remains a critically overlooked failure mode in artificial intelligence for materials science, where systematic patterns of underrepresentation and misrepresentation silently shape which materials are studied, discovered, and deployed while rendering entire classes of materials, synthesis pathways, and knowledge traditions invisible to AI systems. Representational harm is defined here as the systematic underrepresentation, misrepresentation, or exclusion of certain material classes, synthesis methods, research traditions, or communities within materials datasets, resulting in biased AI models that perpetuate inequitable discovery outcomes and reinforce existing power structures in the field. This article articulates five distinct types of representational harm—chemical, structural, synthetic, geographic, and community—along with the four primary mechanisms through which dataset construction choices actively produce these harms, including historical priority, funding asymmetry, measurement accessibility, and publication bias. It further presents a typology of four specific harm failure modes that emerge in materials AI pipelines: invisible classes, distorted property distributions, representational feedback loops, and knowledge colonization. Finally, the paper offers practical detection principles based on diversity, geographic, citation, and community audits as well as five mitigation principles centered on intentional dataset design, data enrichment, weighted representation, inclusion of multiple knowledge systems, and ongoing harm auditing, thereby providing a comprehensive framework for transforming materials dataset construction into a more equitable and epistemically responsible practice.

Keywords Materials informatics, Dataset bias, Representational harm, Materials dataset construction, Epistemic injustice, Fairness in AI

*Correspondence:
Michael Taylor
michael.taylor@outlook.com
[1] Department of AI Materials Systems, University of Auckland, Auckland, New Zealand
[2] Department of Smart Materials Engineering, University of Otago, Dunedin, New Zealand

Introduction

Materials datasets are not neutral. They reflect historical priorities, funding patterns, and community interests in ways that are rarely acknowledged within the technical literature of artificial intelligence for materials science. Some material classes—such as well-characterized oxides or commercially relevant battery chemistries—are massively overrepresented. In contrast, others, including complex intermetallics from low-resource settings or materials developed within non-Western research traditions, are systematically excluded. This pattern of selective inclusion and omission constitutes what this paper identifies as “representational harm,” a failure mode that fundamentally shapes what AI models can learn, what discoveries they can make, and whose scientific questions ultimately receive attention. Rather than treating datasets as passive repositories of objective facts, this analysis treats them as active sites where power relations, epistemic priorities, and historical contingencies are encoded into the very structure of materials knowledge available for machine learning.

The consequences extend far beyond mere statistical imbalance. When AI systems trained on these datasets are deployed for inverse design, property prediction, or autonomous discovery, the representational harm embedded in the training data propagates downstream, limiting the horizon of possible innovations and reinforcing the very imbalances that produced the datasets in the first place. Crawford's seminal work on the atlas of AI powerfully demonstrates how datasets encode power relations at a planetary scale [1], a lesson that has yet to be fully absorbed by the materials informatics community despite its direct relevance to how we curate chemical and structural information. Similarly, Star's early analysis of boundary objects and heterogeneous problem-solving reveals that categorization schemes are never neutral [2]; materials datasets, however, continue to operate under the assumption that their taxonomies are purely technical rather than deeply social and political.

Recent reviews of machine learning applications in molecular and materials science have highlighted impressive advances in predictive accuracy for certain well-studied systems. Yet, they have largely remained silent on the question of whose materials are being modeled and whose are being left behind [3-5]. The Materials Project, for instance, while transformative in accelerating innovation for many classes of inorganic solids, has also become an emblem of how large-scale database construction can inadvertently amplify existing research asymmetries [6, 7]. This paper, therefore, positions representational harm as a core failure mode in materials dataset construction—one that is distinct from, yet entangled with, more commonly discussed technical issues such as data quality or model overfitting. By foregrounding the problem of representational harm, the analysis seeks to shift the discourse in artificial intelligence for materials science from a narrow focus on performance metrics toward a broader consideration of epistemic equity and the long-term societal implications of the datasets we choose to build. The sections that follow define the concept, delineate its types and production mechanisms, present a typology of resulting failure modes, and lay the groundwork for detection and mitigation strategies that can begin to address this pervasive yet under-acknowledged challenge.

Defining Representational Harm

Representational harm arises when the construction of materials datasets systematically underrepresents, misrepresents, or excludes particular material classes, synthesis methods, research traditions, or communities, thereby producing AI systems that embed and amplify these exclusions in downstream discovery processes.

Definition 1: Representational harm is the systematic underrepresentation, misrepresentation, or exclusion of certain material classes, synthesis methods, research traditions, or communities in materials datasets, leading to biased AI systems and inequitable discovery outcomes.

This definition emphasizes that the harm is not accidental or merely statistical but is produced through deliberate and often invisible choices made during data collection, curation, and aggregation. It differs markedly from sampling bias, which is typically understood as a statistical deviation from a presumed population and can, in principle, be corrected through reweighting or resampling techniques. Representational harm, by contrast, concerns the deeper question of which populations are even deemed worthy of inclusion in the first place—what counts as "material" data worth collecting. Measurement error, another frequently discussed issue in materials informatics, refers to inaccuracies in recorded properties or structures for data points that are already present; representational harm operates upstream, determining which data points are allowed to exist at all.

The concept also intersects with, yet remains analytically distinct from, epistemic injustice as articulated in philosophical scholarship and increasingly applied to data practices. While epistemic injustice describes the wrong done to knowers when their contributions are systematically discounted or ignored, representational harm focuses specifically on the material and technical consequences of that discounting within dataset architectures. In materials science, this means that entire research traditions—such as those developed in under-resourced laboratories or drawing on indigenous materials knowledge—may be rendered invisible not because their contributions are judged inaccurate but because they are never invited into the dataset in the first place. As Crawford documents, datasets are never innocent collections of

facts but encode specific visions of what knowledge matters [1]; Star's boundary-object framework further shows how the very act of creating standardized categories for materials data actively excludes alternative ways of knowing [2].

In the context of materials informatics, representational harm manifests as a form of structural exclusion that precedes and enables other forms of bias. When a dataset overrepresents materials studied in high-impact journals from well-funded institutions while neglecting those explored in regional journals or through traditional practices, the resulting AI models inherit and reproduce this skewed worldview. Recent analyses of bias and imbalance in data-driven materials science have begun to acknowledge these issues [8, 9]. Yet, the field still lacks a systematic conceptual framework for naming and addressing representational harm as a distinct failure mode. This section, therefore, establishes that framework, distinguishing representational harm as a foundational epistemic and technical problem that must be confronted before equitable materials AI can be realized.

Figure 1 presents a hierarchical causal architecture linking representational harm types and production mechanisms to dataset distortion, failure modes, epistemic consequences, and corresponding detection and mitigation pathways.

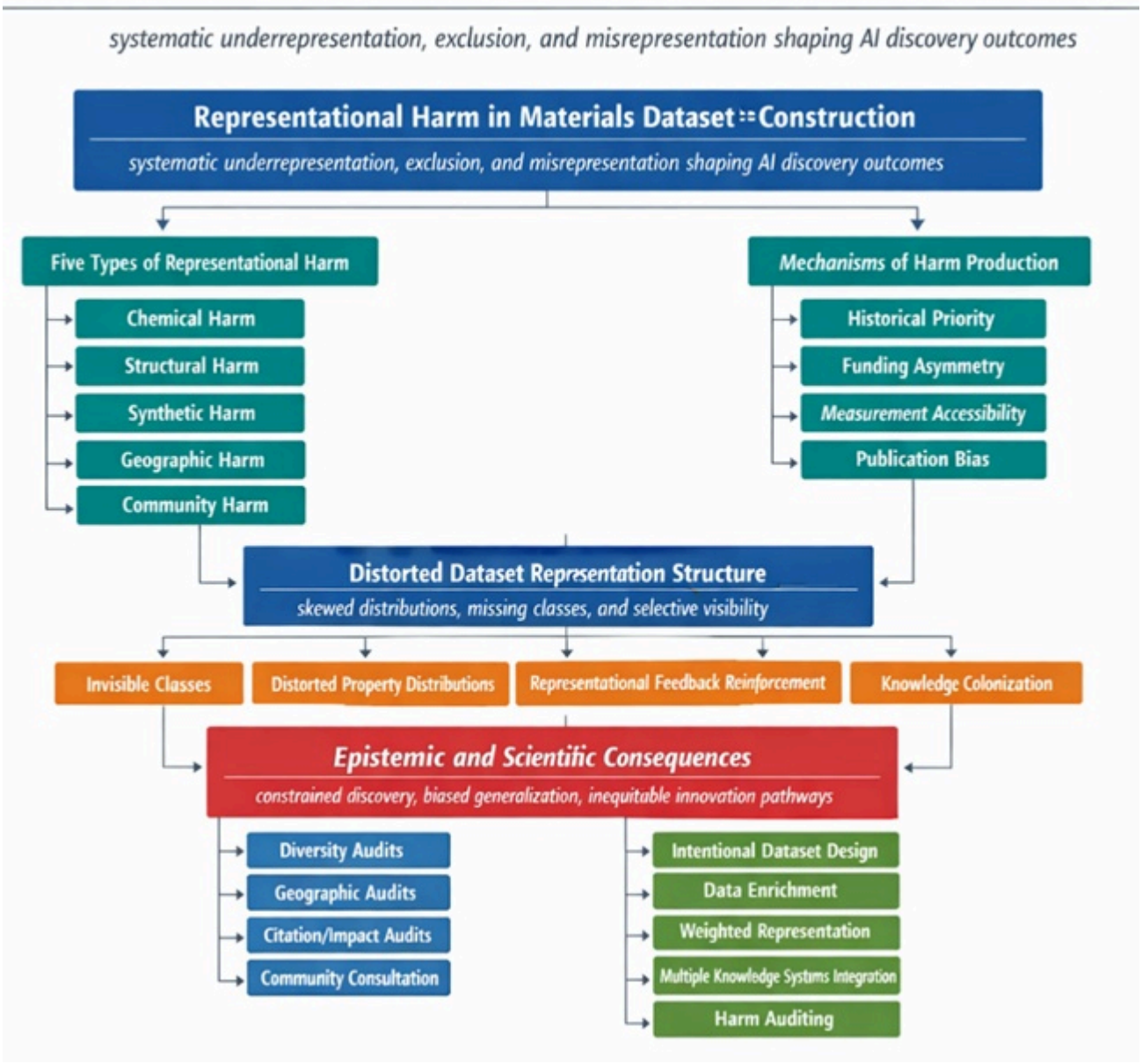

Figure 1. Hierarchical architecture of representational harm in materials dataset construction.

Types of Representational Harm

Five primary types of representational harm can be identified in contemporary materials datasets. Each type operates through distinct but often overlapping mechanisms and produces unique consequences for AI-driven discovery.

Table 1 provides a structural differentiation of representational harm types by linking each to its specific exclusion mechanism, dataset signature, and epistemic risk profile.

Table 1. Structural differentiation between types of representational harm and their epistemic signatures

Harm type	Unit of representation	Primary exclusion mechanism	Observable dataset signature	Epistemic risk produced	Non-redundant diagnostic indicato
Chemical harm	Elements/compositions	Selective study of commercially dominant chemistries	Element frequency skew	Constrained chemical exploration space	Missing plausible periodic table region
Structural harm	Crystal structures/morphology	Preference for ordered systems	Overrepresentation of crystalline phases	Misgeneralization to disordered systems	Absence of amorphous/quasicrys distributions
Synthetic harm	Processing pathways	Bias toward scalable or accessible methods	Limited synthesis diversity	Infeasible AI-recommended pathways	Dominance of a few synthesis technique
Geographic harm	Institutional origin	Concentration in high-resource regions	Geographic clustering of data	Infrastructure-dependent modeling assumptions	Dataset provenance concentration
Community harm	Knowledge systems	Exclusion of non-Western/traditional knowledge	Absence of alternative ontologies	Epistemic homogenization	Missing classificatio diversity

Chemical harm refers to the systematic underrepresentation of specific elements, compositions, or compound families within datasets. For example, while lithium-based battery materials and common perovskites dominate many large-scale repositories, elements such as bismuth, antimony, or rare-earth combinations from developing-world mineral contexts appear far less frequently. The consequence is that AI models trained on these datasets become proficient at optimizing within familiar chemical spaces but fail to generalize to novel or underrepresented chemistries, effectively narrowing the horizon of possible material innovations [10-12].

Structural harm arises from the underrepresentation of certain crystal structures, degrees of disorder, or morphological features. Datasets tend to favor highly ordered, crystalline inorganic solids while undersampling amorphous materials, quasicrystals, or nanostructured composites. This structural skew means that property-prediction models systematically misrepresent the behavior of disordered systems, leading to erroneous extrapolations when AI is asked to design materials with targeted functionalities outside conventional crystallographic norms.

Synthetic harm concerns the underrepresentation of particular synthesis or processing methods. High-throughput computational data often privilege theoretically accessible routes such as solid-state sintering or hydrothermal synthesis, while neglecting mechanochemical, plasma-assisted, or bio-inspired methods that are common in resource-constrained laboratories. As a result, AI systems trained on these datasets recommend synthesis pathways that are inaccessible or impractical for large segments of the global materials research community.

Geographic harm manifests as the overrepresentation of materials data generated in well-funded laboratories located predominantly in wealthy countries. Data from institutions in North America, Western Europe, and parts of East Asia crowd out contributions from laboratories in the Global South, even when those contributions address locally relevant materials challenges such as sustainable extraction or climate-resilient ceramics. This geographic skew not only distorts the global map of materials knowledge but also embeds assumptions about infrastructure and resource availability that do not hold universally.

Type 5: Community harm involves the exclusion of traditional, indigenous, or non-Western material knowledge systems. Datasets built primarily on peer-reviewed Western scientific literature rarely incorporate ethnobotanical knowledge of natural composites, artisanal metallurgy practices, or community-based material innovations developed outside formal academic channels. The erasure of these knowledge systems constitutes a profound form of representational harm that severs AI from potentially transformative sources of insight while reinforcing colonial patterns of knowledge extraction [13-16].

These five types are not mutually exclusive; a single dataset can exhibit multiple forms of harm simultaneously, amplifying their collective impact on materials AI.

Mechanisms of Harm Production

Representational harm is not an inevitable byproduct of data collection but is actively produced through four interlocking mechanisms that operate at the level of dataset construction itself.

Historical Priority operates when materials that were intensively studied in earlier decades continue to dominate contemporary datasets simply because legacy data are easier to digitize and aggregate. Once a compound class enters a major database, subsequent curation efforts naturally build upon it, creating a self-reinforcing cycle of overrepresentation.

Funding Asymmetry arises because well-funded research programs generate far more publishable data points than under-resourced ones. Grant-winning laboratories in high-income countries produce high-volume, standardized measurements that flow readily into open repositories. In contrast, laboratories in lower-resource settings often lack the infrastructure to contribute at a comparable scale.

Measurement Accessibility privileges properties and materials that are straightforward to characterize with widely available laboratory equipment. Easily measured band gaps or formation energies appear in abundance, whereas exotic or hazardous properties requiring specialized facilities remain underrepresented, skewing the overall distribution of knowledge encoded in datasets [17-20].

Publication Bias ensures that only positive or “successful” results enter the literature and, by extension, the datasets derived from it. Negative results, failed syntheses, and null findings—particularly those involving underrepresented material classes—are rarely published, creating systematic gaps that AI models interpret as absence rather than unexplored possibilities.

These mechanisms interact dynamically. Historical priority is amplified by funding asymmetry, which in turn favors measurement-accessible systems, all of which are filtered through publication bias. The result is a dataset architecture

that appears technically neutral yet is deeply shaped by social, economic, and historical contingencies.

Table 2 establishes a cross-level mapping between harm production mechanisms, emergent failure modes, and targeted detection and mitigation interventions.

Table 2. Mapping representational harm mechanisms to failure modes and targeted intervention levers

Harm mechanism	Dataset-level effect	Failure mode triggered	Downstream AI behavior	Detection lever	Mitigation lever
Historical priority	Legacy data dominance	Invisible classes	Reproduction of historical focus areas	Diversity audit	Intentional dataset design
Funding asymmetry	Data volume imbalance	Representational Feedback Reinforcement	Over-optimization of dominant classes	Geographic audit	Data enrichment
Measurement accessibility	Property measurement bias	Distorted Property Distributions	Overconfident extrapolation	Citation/impact audit	Weighted representation
Publication bias	Absence of negative results	Invisible Classes + Distortion	False absence assumptions	Diversity audit	Harm auditing
Cross-mechanism interaction	Compounded skew	Knowledge Colonization	Suppression of alternative knowledge systems	Community consultation	Multiple knowledge systems integration

A Typology of Harm Failure Modes

Building upon the types and mechanisms identified above, this section presents a typology of four distinct harm failure modes that emerge when representational harm propagates through materials AI pipelines.

Invisible classes occur when entire material classes are absent from training datasets, rendering them literally undiscoverable by AI systems. The mechanism is straightforward exclusion during construction; the detection signature is a zero or near-zero representation of a chemically plausible class despite its documented existence in the broader literature. Distorted property distributions arise when the limited samples that do exist for a given class fail to capture the true range of possible properties, leading models to extrapolate poorly or confidently predict impossible behaviors. For instance, if only high-stability perovskites are present, the model may systematically underestimate the stability window of novel variants. The representational feedback loop constitutes the most insidious failure mode. AI models trained on biased datasets recommend new experiments that preferentially target already overrepresented classes; those new data are then fed back into the dataset, further entrenching the original imbalance. Conceptually, the representational harm cycle can be visualized as a closed loop in which initial dataset construction biases feed into AI model training, which then influences discovery outcomes that in turn drive targeted dataset augmentation back toward the same biased classes, creating a self-perpetuating cycle of representational reinforcement. Knowledge Colonization emerges when dominant datasets overwrite or marginalize alternative knowledge systems. Once a global materials AI platform trained on Western-centric data is adopted worldwide, local research traditions risk being reframed as “validation cases” rather than primary sources, effectively colonizing epistemic space [21-24].

Each of these failure modes represents a distinct pathway through which representational harm undermines the reliability, equity, and epistemic integrity of materials AI.

Detection Principles

Detecting representational harm in materials datasets requires moving beyond conventional performance metrics to systematic, multi-layered auditing practices that expose the epistemic and structural exclusions embedded in data construction. Four core detection principles provide a practical foundation for identifying these harms before they propagate into AI pipelines.

Diversity audits involve quantitative and qualitative assessments of representation across chemical, structural, and synthetic dimensions. Rather than simply counting total entries, diversity audits map the distribution of material classes against a baseline of scientific plausibility derived from broader literature surveys, revealing systematic absences that statistical summaries might otherwise mask. For instance, a diversity audit might quantify the proportion of entries containing underrepresented elements such as those prevalent in tropical or arid-region mineralogies, thereby surfacing chemical harm that would remain invisible in aggregate accuracy scores [25-27].

Geographic audits track the provenance of data points to expose imbalances in laboratory origins and funding ecosystems. By geocoding contributor institutions and cross-referencing with national research expenditure data, these audits illuminate how well-resourced regions dominate the knowledge base. At the same time, contributions from lower-resource settings remain marginal. Geographic audits thus operationalize the detection of geographic harm, making visible the ways in which dataset construction mirrors global inequalities in scientific infrastructure.

Citation/impact audits compare dataset representation against independent measures of scientific importance, such as citation networks or patent landscapes, to identify mismatches between a material class's documented research value and its presence in training data. When high-impact but underrepresented classes—such as those emerging from community-driven sustainability initiatives—appear at disproportionately low frequencies, the audit flags community harm and publication bias in action.

Community consultation engages directly with affected research traditions and local practitioner communities to validate or challenge the dataset's representational choices. This principle recognizes that quantitative audits alone cannot capture epistemic exclusions rooted in non-Western or indigenous knowledge systems; instead, structured dialogues with knowledge holders reveal whether traditional synthesis routes or material classifications have been rendered invisible. As Star's boundary-object framework reminds us, categories are never neutral [2], and community consultation ensures that the lived experience of underrepresented groups informs the detection process.

Taken together, these detection principles form an integrated protocol that shifts the burden of proof onto dataset curators, requiring them to demonstrate not only technical completeness but epistemic representativeness. Recent scholarship on bias and imbalance in data-driven materials science underscores the urgency of such audits [9]. At the same time, Crawford's analysis of how datasets encode power relations [1] further justifies their routine application. Without proactive detection grounded in these principles, representational harm will continue to operate undetected, silently constraining the future of materials discovery. Implementing these audits early in the dataset lifecycle allows the field to move from reactive correction to preventive equity, ensuring that materials AI systems are built on foundations that reflect the full diversity of global materials knowledge.

Mitigation Principles

Mitigation of representational harm demands proactive, design-level interventions that reorient dataset construction from passive aggregation toward deliberate epistemic inclusion. Five interlocking principles offer an actionable framework for

reducing harm while preserving the technical integrity required for robust AI modeling.

Intentional dataset design begins at the earliest planning stage by explicitly defining target representation thresholds for each material class, synthesis method, and geographic origin. Curators set minimum inclusion quotas derived from scientific importance metrics rather than historical availability, thereby countering historical priority and funding asymmetry before data collection even begins.

Data enrichment requires active outreach to underrepresented sources, including gray literature from regional journals, laboratory notebooks from low-resource settings, and oral or practice-based knowledge from traditional communities. This principle transforms curation from a harvesting exercise into a generative one, deliberately seeking and digitizing data that would otherwise remain outside mainstream repositories [24-29].

Weighted representation techniques during dataset assembly and training involve oversampling underrepresented classes through synthetic augmentation, transfer learning from related but better-sampled domains, or stratified resampling protocols. By adjusting the effective frequency of neglected material classes, weighted representation corrects distorted property distributions without fabricating false data points.

Multiple knowledge Systems integration explicitly incorporates alternative classification schemes and property ontologies drawn from non-Western research traditions alongside conventional Western taxonomies. This principle acknowledges that material knowledge is plural and prevents knowledge colonization by ensuring that dominant datasets do not overwrite parallel epistemic frameworks.

Harm auditing establishes recurring, mandatory reviews of datasets using the detection principles outlined previously, with public documentation of representational metrics and corrective actions taken. Regular harm auditing institutionalizes accountability, ensuring that mitigation remains an ongoing practice rather than a one-time intervention.

These mitigation principles are mutually reinforcing. Intentional design sets the strategic direction, data enrichment populates the gaps, weighted representation balances the distributions, multiple knowledge systems expand the epistemic scope, and harm auditing sustains the entire process over time. Butler and colleagues have shown how machine learning can accelerate materials science when applied thoughtfully [4], yet without these mitigation steps, acceleration risks amplify existing exclusions. Similarly, the seed reference on representational harm itself calls for precisely this kind of systematic response within materials informatics [3]. By embedding these principles into standard operating procedures for major materials databases, the community can begin to dismantle the mechanisms of harm production identified earlier and interrupt the feedback loops that currently perpetuate representational failure modes. The result is not merely fairer datasets but more scientifically robust ones, capable of discovering materials that current systems are structurally blind to. Implementation will require new funding streams, revised peer-review criteria, and cross-disciplinary collaboration between materials scientists, data ethicists, and community representatives. Yet, the long-term payoff is a genuinely global materials AI enterprise.

Relation to Other Failure Modes

Representational harm does not exist in isolation but intersects with, amplifies, and sometimes masquerades as several other well-documented failure modes in materials informatics. Understanding these relations clarifies both the distinctiveness of representational harm and the necessity of addressing it as a foundational concern.

First, representational harm frequently manifests as a deeper cause of sampling bias, which is conventionally treated as a purely statistical issue amenable to reweighting. While sampling bias concerns deviations from an assumed population distribution, representational harm determines which population is even considered in the first place. A dataset that systematically excludes entire chemical families does not merely suffer from imbalance; it has already predefined an artificially narrow population, rendering subsequent statistical corrections superficial.

Second, representational harm provides the concrete technical substrate for epistemic injustice in materials science. Where epistemic injustice describes the philosophical wrong of discounting certain knowers, representational harm translates that wrong into dataset architecture: entire research traditions become unmodelable because their data were never collected. The connection is direct—epistemic injustice at the level of knowledge production becomes representational harm at the level of data curation.

Third, representational harm functions as an upstream driver of algorithmic bias in deployed models. When training data encode geographic or community exclusions, the resulting AI systems inherit and propagate those exclusions, producing predictions that systematically favor overrepresented classes. Algorithmic bias is therefore not an independent failure but the downstream consequence of unaddressed representational harm.

Finally, representational harm operationalizes data colonialism within materials AI. Crawford's analysis of how AI datasets encode planetary power relations [1] finds a precise parallel in the way dominant materials databases extract value from global scientific labor while marginalizing contributions from the Global South and indigenous communities. Data colonialism is not merely a metaphor here; it describes the concrete process by which knowledge produced in one context is appropriated and recentered within Western-centric repositories.

By mapping these relations, the analysis reveals representational harm as a linchpin failure mode—one that must be confronted if progress on sampling bias, epistemic injustice, algorithmic bias, or data colonialism is to be more than cosmetic. Addressing it requires the integrated detection and mitigation principles developed above, ensuring that materials AI advances equity rather than entrenching existing hierarchies.

Implications for Materials AI Practice

The recognition of representational harm as a core failure mode carries immediate and far-reaching implications for every stakeholder involved in materials dataset construction and AI model development.

For dataset creators, three concrete changes are required: first, every new or updated dataset must include a mandatory representational harm audit using the principles outlined in Section 6; second, curation documentation must explicitly report representation metrics across the five harm types; and third, active mitigation strategies must be implemented and publicly justified before dataset release.

For model developers, practice must shift from treating training data as given to treating it as a site of ethical and epistemic responsibility. Developers should routinely assess representational harms in source datasets, test model performance on deliberately diverse holdout sets that include underrepresented classes, and document how training choices interact with known harm mechanisms.

For reviewers and journal editors, evaluation criteria must expand to include scrutiny of dataset representation. Reviewers should ask whether authors have conducted diversity and geographic audits, whether mitigation steps were taken, and whether the paper acknowledges potential representational harms rather than assuming dataset neutrality.

For the broader materials AI community, two systemic actions are essential: first, the development of shared representation standards and metadata schemas that make representational metrics machine-readable and comparable across repositories; and second, targeted funding initiatives specifically dedicated to data collection in neglected material classes, geographic regions, and knowledge systems.

These implications collectively transform materials AI from a technically focused enterprise into one that is also epistemically accountable. As Schmidt and colleagues have highlighted in their review of machine learning in solid-state materials science [5], the field stands at a crossroads where continued growth without equity risks producing powerful but

narrowly applicable tools. By embedding the principles and frameworks developed here into everyday practice, the community can ensure that future advances in artificial intelligence for materials science serve the full diversity of global materials challenges rather than a privileged subset.

Conclusion

Representational harm in materials dataset construction constitutes a fundamental yet largely unacknowledged failure mode in artificial intelligence for materials science. This paper has defined the concept, delineated its five primary types and four production mechanisms, presented a typology of four resulting harm failure modes, and offered concrete detection and mitigation principles grounded in epistemic equity. By tracing the pathway from dataset construction choices to downstream discovery limitations, the analysis demonstrates that representational harm is not peripheral but central to the reliability, fairness, and societal value of materials AI systems.

The representational harm cycle—visualized in the closed loop of biased construction, model training, discovery outcomes, and further dataset reinforcement—will persist unless the field adopts the deliberate, multi-principle approach advocated here. Crawford's warning that datasets encode power relations, combined with Star's insight into the non-neutrality of categories, provides a compelling mandate for change. The future of materials discovery depends not only on larger models or more data but on datasets that genuinely reflect the planetary diversity of materials knowledge.

This failure mode analysis, therefore, issues a clear call: representational harm awareness must become a standard component of materials dataset construction. Only through intentional, audited, and inclusive practices can the materials informatics community realize the full promise of artificial intelligence while avoiding the reproduction of historical and structural inequities. The framework presented offers a practical path forward—one that transforms dataset curation from an invisible infrastructure into a site of epistemic justice and scientific creativity.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 28 Feb 2025 Revised: 13 Apr 2025 Accepted: 26 May 2025
Published online: 18 July 2025

References

Şaan T. The atlas of AI: Power, politics, and the planetary costs of artificial intelligence. J Consum Consum Res. 2024;16(1):181-7.

Harvey F. Heterogeneous distributed problem-solving involving visual objects as boundary objects. Front Commun. 2024;8:1275695.

Wang A, Barocas S, Laird K, Wallach H. Measuring representational harms in image captioning. In: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency; 2022. p. 324-35.

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.

Li R, Zhao W, Li R, Gan C, Chen L, Wang Z, et al. Leveraging machine learning for accelerated materials innovation in lithium-ion battery: A review. J Energy Chem. 2025;106:44-62.

Fujinuma N, DeCost B, Hattrick-Simpers J, Lofland SE. Why big data and compute are not necessarily the path to big materials science. Commun Mater. 2022;3(1):59.

De Breuck PP, Evans ML, Rignanese GM. Robust model benchmarking and bias-imbalance in data-driven materials science: A case study on modnet. J Phys Condens Matter. 2021;33(40):404002.

DeCost BL, Hattrick-Simpers JR, Trautt Z, Kusne AG, Campo E, Green ML. Scientific AI in materials science: A path to a sustainable and scalable paradigm. Mach Learn Sci Technol. 2020;1(3):033001.

Dai K, Kim J, Džeroski S, Wicker J, Dobbie G, Dost K. Assessing the risk of discriminatory bias in classification datasets. Mach Learn. 2025;114(9):204.

Mannodi-Kanakkithodi A, McDannald A, Sun S, Desai S, Brown KA, Kusne AG. A framework for materials informatics education through workshops. MRS Bull. 2023;48(5):560-9.

Pyzer-Knapp EO, Manica M, Staar P, Morin L, Ruch P, Laino T, et al. Foundation models for materials discovery–current state and future directions. npj Comput Mater. 2025;11(1):61.

Gibaldi M, Kapeliukha A, White A, Luo J, Mayo RA, Burner J, et al. MOSAEC-DB: A comprehensive database of experimental metal–organic frameworks with verified chemical accuracy suitable for molecular simulations. Chem Sci. 2025;16(9):4085-100.

Issa R, Hamad S, Grau-Crespo R, Awad E, Sparks TD. Sample-efficient active learning for materials informatics using integrated posterior variance. Comput Mater Sci. 2026;267:114551.

Herzog C, Branford J. Relational ethics and structural epistemic injustice of AI in medicine. Philos Technol. 2025;38(4):1-28.

Chakraborty S, Björk J, Dahlqvist M, Rosen J, Heintz F. A survey of AI-supported materials informatics. Comput Sci Rev. 2026;59:100845.

Liu T, Tho ZY, Barnard AS. Understanding the importance of individual samples and their effects on materials data using explainable artificial intelligence. Digit Discov. 2024;3(2):422-35.

Otyepka M, Pykal M, Otyepka M. Advancing materials discovery through artificial intelligence. Appl Mater Today. 2025;47:102981.

Liu Y, Li S, Yang B, Liu Y, Deng Y, Yu J, et al. Materials dataset from microgravity levitation experiments on the China Space Station. Sci Data. 2026;13:112.
https://doi.org/10.1038/s41597-025-06428-0.

Huerta EA, Blaiszik B, Brinson LC, Bouchard KE, Diaz D, Doglioni C, et al. FAIR for AI: An interdisciplinary and international community building perspective. Sci Data. 2023;10(1):487.

Trezza G, Chiavazzo E. Classification-based detection and quantification of cross-domain data bias in materials discovery. J Chem Inf Model. 2024;65(4):1747-61.

Bennett SJ, Claisse C, Luger E, Durrant AC. Unpicking epistemic injustices in digital health: On the implications of designing data-driven technologies for the management of long-term conditions. In: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society; 2023. p. 322-32.

Pieper M. Is data material? Toward an environmental sociology of AI. AI Soc. 2025:1-2.

Wang L, Kameswaran V, Kacorri H. Toward a taxonomy of algorithmic harms for disability: A systematic review. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. 2025;8(3):2649-65.

Sarhan H, Hegelich S. Understanding and evaluating harms of AI-generated image captions in political images. Front Polit Sci. 2023;5:1245684.

Choi I, Zu J. Generating language assessment content free from representational harms. Lang Test. 2025;42(4):539-60.

Ahmad A, Vallès Y, Idaghdour Y. Bias in AI systems: Integrating formal and socio-technical approaches. Front Big Data. 2025;8:1686452.

Tempke RS. Artificial intelligence based approach for rapid material discovery: From chemical synthesis to quantum materials [doctoral dissertation]. West Virginia: West Virginia University.