

ORIGINAL RESEARCH

Open access

Algorithmic Path Dependency as Scientific Path Dependency: A Conceptual Link

Ahmed Benali^{1*}, Karim Boudiaf², Samir Touati¹

Abstract

In the rapidly evolving domain of artificial intelligence for materials science, path dependency remains a critically overlooked phenomenon that shapes both computational pipelines and the broader scientific enterprise. Algorithmic path dependency manifests when seemingly innocuous early choices in neural network initialization, training data ordering, hyperparameter selection, feature descriptor definition, or early stopping criteria create irreversible constraints on subsequent model behaviors and outputs, as evidenced in recurrent neural network architectures designed for heterogeneous materials. Scientific path dependency, by contrast, arises in the history and philosophy of science when initial decisions regarding research problems, material systems, theoretical frameworks, experimental protocols, or funding priorities lock research communities into particular trajectories, rendering alternative avenues increasingly difficult to pursue even when they might yield superior insights. This paper advances the theoretical claim that algorithmic path dependency propagates directly into scientific path dependency within materials AI, such that technical decisions made at the level of code and data become de facto determinants of which materials are discovered, which questions are asked, and which knowledge ultimately enters the scientific canon. The linkage operates through identifiable mechanisms, including output filtering, resource allocation, knowledge representation, and publication bias, each amplifying the long-term scientific consequences of early algorithmic commitments. By drawing upon foundational economic concepts of increasing returns and historical contingency alongside contemporary literature in machine learning for materials, this theoretical analysis proposes that materials AI researchers must explicitly recognize these dynamics to avoid unintended lock-in effects that could limit the diversity and robustness of future discoveries. The analysis further derives corollaries concerning constrained output diversity, the practical irreversibility of certain scientific paths, and the necessity of methodological pluralism, offering concrete implications for research practice, peer review standards, and community norms. Ultimately, this conceptual linkage reframes early algorithmic decisions not as mere technical details but as foundational scientific commitments whose consequences reverberate through the entire materials discovery ecosystem.

Keywords Materials AI, Machine learning, Path dependency, Algorithmic path dependency, Scientific path dependency, Historical contingency

*Correspondence:

Ahmed Benali
ahmed.benali@gmail.com

¹ Department of Intelligent Materials Systems, University of Algiers, Algiers, Algeria

² Department of Materials Data Engineering, University of Tunis El Manar, Tunis, Tunisia

Introduction

Materials AI has transformed the landscape of materials discovery by enabling rapid screening of vast compositional spaces, prediction of novel structures, and guidance of

experimental efforts toward promising candidates [1-5]. Yet beneath the surface of these impressive capabilities lies a fundamental issue that has received insufficient theoretical attention: the role of path dependency. When researchers select a particular dataset, initialize a neural network with

random weights, choose specific hyperparameters, define material descriptors in one way rather than another, or apply early stopping rules during training, these decisions are frequently treated as purely technical or arbitrary matters of implementation convenience [6-11]. In reality, however, such choices establish trajectories that become self-reinforcing and increasingly resistant to reversal, even when subsequent evidence suggests superior alternatives might exist.

This paper argues that algorithmic path dependency—the phenomenon whereby early computational choices constrain future algorithmic possibilities—directly translates into scientific path dependency, whereby early research decisions constrain future scientific directions in materials discovery. The problem is not merely that different initializations might produce slightly different models; rather, the very questions scientists can ask, the materials they choose to investigate experimentally, and the knowledge they ultimately generate are shaped by these upstream algorithmic commitments [3, 12-14]. For instance, if an active learning pipeline prioritizes certain compositional regions based on an initial random seed or data ordering, entire families of materials may remain unexplored, not because they lack promise but because they were never surfaced by the algorithm [7].

The consequences of ignoring this linkage are profound. Materials AI is increasingly positioned as an autonomous or semi-autonomous driver of discovery [7], yet without a conceptual framework for understanding how code-level decisions become science-level constraints, the field risks locking itself into narrow technological and epistemic paths. Early decisions in model architecture or descriptor choice do not simply affect predictive accuracy in isolation; they determine which phenomena can be represented, which hypotheses can be tested, and which breakthroughs can be recognized as such [13, 15-18]. This theoretical analysis, therefore, seeks to make path dependency visible as a core feature rather than a peripheral bug in materials AI systems.

By tracing the origins of path dependency concepts from economics and the history of technology [1, 2], defining algorithmic and scientific variants with precision, establishing their linkage through explicit mechanisms, deriving logical corollaries, and outlining practical implications, the present work offers a conceptual scaffold for more reflexive practice in the field. The analysis proceeds without reliance on new empirical data,

simulations, or performance metrics, focusing instead on logical elaboration, conceptual distinctions, and illustrative examples drawn from the existing literature. In doing so, it proposes that seemingly neutral technical choices in AI pipelines carry scientific weight comparable to the selection of research questions or theoretical paradigms [3, 19-23]. Recognition of this reality is essential if materials AI is to fulfill its promise of accelerating discovery without inadvertently narrowing the scope of what can be discovered.

Path Dependency: Origins and Concepts

Path dependency has long been recognized as a fundamental property of complex dynamic systems in which history matters in a non-trivial way.

Path dependency is a property of processes in which early events or decisions disproportionately constrain subsequent outcomes, creating forms of lock-in that persist even when superior alternatives become available later.

This definition draws directly from foundational work in economics and the history of technology. Arthur demonstrated how increasing returns—situations in which the benefits of adopting a technology or standard grow with the number of adopters—can lead to lock-in effects where an initially dominant but ultimately inferior option becomes entrenched [1]. Similarly, David illustrated the persistence of the QWERTY keyboard layout despite the existence of ergonomically superior alternatives, showing how small historical contingencies combined with switching costs can produce long-term inefficiency [2].

Key concepts associated with path dependency include increasing returns, switching costs, lock-in, and historical contingency. Increasing returns refer to positive feedback loops whereby the adoption of one option makes its continued dominance more likely. Switching costs encompass the economic, cognitive, or infrastructural expenses required to change paths once established. Lock-in describes the resulting state of rigidity in which change becomes difficult or impossible without external disruption. Historical contingency emphasizes that outcomes depend not only on current conditions but on the specific sequence of past events, rendering the process non-ergodic—different sequences can lead to different equilibria even under identical starting conditions.

These concepts have proven generative beyond economics, finding application in technology studies, institutional theory, and, more recently, in computational domains. In the present context, they provide a lens for understanding why early choices in materials AI pipelines can exert outsized influence on downstream scientific outcomes. The non-ergodic nature of path-dependent processes implies that two otherwise identical research programs employing different initializations or data orderings may converge on fundamentally different material discoveries and theoretical understandings, even when both appear equally “optimal” by conventional metrics [3].

Importantly, path dependency is not synonymous with mere inertia or randomness. It involves genuine sensitivity to initial conditions coupled with self-reinforcing mechanisms that amplify small differences over time. This distinguishes it from stochastic variation that averages out across multiple runs. In path-dependent systems, the amplification is structural: early choices reshape the possibility space itself, foreclosing certain futures while enabling others [1]. The remainder of this analysis applies these foundational ideas to algorithmic and scientific domains within materials AI, demonstrating how economic and historical concepts illuminate contemporary computational practices in materials discovery.

Algorithmic Path Dependency

Building upon the general concept of path dependency, algorithmic path dependency refers specifically to the constraining effects of early computational choices within machine learning pipelines.

Algorithmic path dependency occurs when initial decisions regarding model architecture, parameter initialization, data sequencing, hyperparameter configuration, or training termination criteria create self-reinforcing trajectories that limit the set of reachable model states and outputs, rendering later corrections costly or infeasible.

At least five primary sources contribute to algorithmic path dependency in materials AI contexts.

Random initialization of neural network weights. Even identical architectures trained on the same data can converge to qualitatively different internal representations when initialized differently, particularly in non-convex

optimization landscapes typical of deep learning for materials properties [8, 9].

Ordering of training data. The sequence in which examples are presented during stochastic gradient descent influences gradient trajectories and final model parameters; in materials datasets characterized by compositional imbalance, early exposure to certain chemistries can bias the learned embedding space [12, 18].

Choice of hyperparameters. Learning rate schedules, regularization strengths, or batch sizes selected at the outset establish optimization dynamics that become difficult to escape without full retraining, effectively locking the model into particular generalization behaviors [10, 11].

Selection of features or descriptors. Once a specific set of material descriptors or graph representations is chosen, the model's inductive bias is fixed; alternative physical or chemical encodings that might reveal different structure-property relationships are structurally excluded from consideration [5, 13].

Early stopping or convergence criteria. Decisions about when to terminate training based on validation metrics made early in development can prevent exploration of flatter minima or more generalizable solutions that emerge only after prolonged optimization [15, 24].

These sources interact multiplicatively. For example, a particular random initialization combined with a specific data ordering can amplify sensitivity to hyperparameter choices, creating compound lock-in effects observable in recurrent neural networks applied to path-dependent mechanical responses in heterogeneous materials [8, 25-28]. The consequences are not merely quantitative differences in accuracy but qualitative differences in the model's representational capacity: certain material phenomena may be accurately captured. In contrast, others remain invisible or poorly modeled [12].

Literature on physically recurrent neural networks for path-dependent materials explicitly acknowledges sensitivity to initialization and training protocols, demonstrating that different starting conditions yield distinct predictions of history-dependent behavior even when loss functions are identical [8, 28]. Similarly, deep learning approaches to plasticity prediction reveal that data ordering profoundly affects the learned constitutive relations, with early exposure to certain strain paths locking the network into

specific deformation regimes [12]. These examples illustrate how algorithmic path dependency operates as a structural feature rather than a transient noise that can be averaged away through ensemble methods alone. Once established, the path constrains not only predictive performance but the very hypotheses the model can support when deployed in discovery workflows [3].

Scientific Path Dependency

Scientific path dependency extends the logic of historical contingency to the epistemic practices of research communities.

Scientific path dependency arises when initial choices of research problems, material systems, theoretical frameworks, experimental techniques, or publication and funding priorities create self-reinforcing trajectories that constrain the questions asked, the evidence considered, and the knowledge ultimately produced, even when alternative paths might prove more fruitful.

At least five primary sources fuel scientific path dependency in materials science and related fields.

Choice of research problem. Early selection of particular phenomena or material classes directs collective attention and resources, making subsequent shifts to unrelated systems cognitively and institutionally expensive [19, 23].

Selection of materials to study. Focusing experimental or computational efforts on a narrow compositional space—often guided by initial algorithmic suggestions—renders broader exploration less likely over time [7, 14].

Once a community standardizes around specific characterization methods or synthesis routes, alternative techniques become marginalized, limiting the types of data that enter the evidentiary base [20, 24, 25].

Theoretical frameworks employed. Adoption of particular modeling paradigms or descriptor sets early on shapes the conceptual vocabulary and explanatory strategies available to subsequent researchers [6, 13].

Publication and funding decisions. Successful initial results in one direction generate citations, grants, and visibility that reinforce the same trajectory while alternative approaches remain unpublished or unfunded [19, 23].

These sources interact through feedback loops between community practices and individual research programs. Historical contingency studies in philosophy of science emphasize how small early differences in problem framing can lead to divergent research programs that appear incommensurable years later [23]. In materials AI specifically, the integration of machine learning into discovery pipelines introduces new vectors for such contingency: algorithmic recommendations that prioritize certain candidates shape which materials receive experimental validation, which in turn influences future funding calls and publication norms [7, 14].

Literature on the history of artificial intelligence highlights parallel dynamics in which early architectural choices and benchmark selections have locked entire subfields into particular evaluation criteria and research questions [19]. Analogous contingency effects appear in ecological and evolutionary studies where historical sequence determines community assembly outcomes [20, 21, 25]. Within materials contexts, these patterns manifest when initial algorithmic outputs bias the selection of systems for inverse design, thereby constraining the functional properties ultimately targeted [6]. The result is a form of epistemic lock-in in which the scientific community's collective knowledge trajectory reflects not only intrinsic material realities but also the path-dependent history of its algorithmic tools [3].

Theoretical Claim: Algorithmic and Scientific Path Dependency Are Linked

The core theoretical contribution of this analysis is the explicit linkage between algorithmic and scientific forms of path dependency in the specific context of materials AI.

In materials AI, algorithmic path dependency propagates into scientific path dependency such that early computational choices constrain not only model outputs but also the scientific questions that can be asked, the materials that can be discovered, and the knowledge that can be generated.

The linkage between algorithmic and scientific path dependency occurs through multiple reinforcing mechanisms (elaborated in subsequent sections). It produces long-term scientific consequences that are not easily reversed once established, rendering seemingly

technical algorithmic decisions into de facto scientific commitments.

This theoretical linkage reframes materials AI not as a neutral tool but as an active participant in shaping scientific possibility spaces.

Table 1 distinguishes algorithmic from scientific path dependency across locus, mechanisms, switching costs, and forms of lock-in, clarifying why the former can systematically propagate into the latter in materials AI.

Table 1. Analytical Comparison of Algorithmic and Scientific Path Dependency in Materials AI

Dimension	Algorithmic path dependency	Scientific path dependency	Co-significance
Primary locus	Model training pipeline and computational workflow	Research community, laboratory practice, and disciplinary agenda	Essential to the existing technical ecosystem
Typical early choices	Initialization, data ordering, hyperparameters, descriptor selection, and stopping criteria	Problem framing, material selection, technique adoption, theoretical framing, and publication/funding priorities	Stable and broadly accepted
Immediate effect	Narrow reachable model states and output distributions	Narrows acceptable questions, candidate systems, and evidentiary standards	Classical reductionist optimization

Self-reinforcing mechanism	Optimization lock-in, representational inertia, reuse of trained pipelines, and benchmark dependence	Citation accumulation, grant reinforcement, experimental replication, and institutional legitimacy	Critique and resistance
Switching costs	Retraining, redesign, data reprocessing, loss of comparability, and engineering overhead	Reframing projects, rebuilding infrastructure, retraining personnel, reputational and funding risk	Experimental differentiation, which allows for alternative paths
Observable manifestation	Persistent output sensitivity, constrained candidate ranking, and stable but narrow search behavior	Repeated focus on the same material classes, methods, and explanatory vocabularies	Material dependency, entrenched research paradigms
Unit of lock-in	Model behavior and representational capacity	Research trajectory and knowledge canon	Disciplinary boundaries, rigidity
Temporal horizon	Emerges during development and deployment cycles	Accumulates across projects, publications, and field-level agenda-setting	Short-term success, also contributing to long-term lock-in
Main risk	Invisible exclusion of alternative outputs	Invisible exclusion of alternative discoveries and theories	Shallow materialism, narrow view
Preferred mitigation	Multi-seed training, alternative descriptors, reordered datasets, parallel	Methodological pluralism, broader evaluation standards, and diversified	Diversification and integration

	models, and sensitivity analysis	funding/publication incentives
--	----------------------------------	--------------------------------

Figure 1 conceptualizes how early algorithmic commitments propagate through four mediating mechanisms to generate downstream scientific path dependency in materials AI.

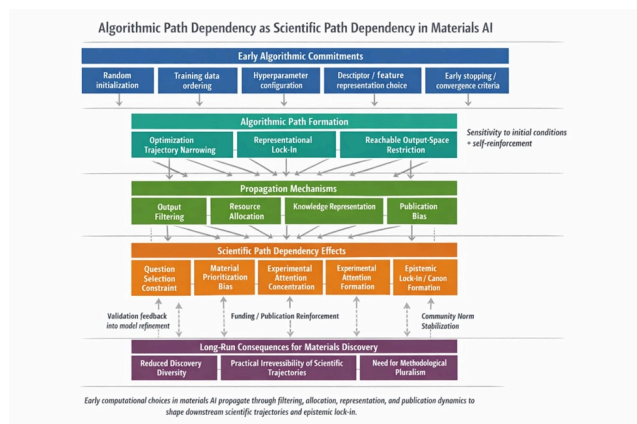


Figure 1. Algorithmic path dependency as scientific path dependency in materials AI

Algorithmic choices do not merely assist discovery; they co-determine the epistemic horizon of the field. For example, when a Bayesian active learning framework initialized with particular priors surfaces only certain compositional regions [14], the resulting experimental validations reinforce those same regions in future iterations, creating a feedback loop between code and laboratory practice. The seed paper on this exact conceptual linkage already anticipates that such propagation occurs systematically in materials AI pipelines [3].

Conceptually, the linkage can be visualized as a directed flow diagram in which algorithmic decision nodes at the top (initialization, data ordering, descriptor choice) feed downward through filtering, allocation, representation, and publication mechanisms into scientific constraint nodes at the bottom (question selection, material prioritization, knowledge canon formation). Arrows between nodes thicken over time to represent increasing returns and lock-in, while feedback loops ascend from scientific outcomes back to algorithmic refinement, illustrating the co-evolutionary nature of the dependency. This conceptual figure underscores that the propagation is bidirectional yet asymmetrically initiated by algorithmic choices in contemporary practice.

The claim rests on logical entailment rather than empirical induction: given the definitions established in prior sections and the documented sensitivity of both algorithmic and scientific processes to early conditions [1, 2, 8, 23], the propagation follows necessarily once materials AI is positioned as the primary interface between computation and discovery [4, 7]. Because algorithmic representations become the lens through which materials are perceived and prioritized, constraints at the algorithmic level become epistemic constraints at the scientific level [5, 13]. This perspective offers a novel theoretical bridge between computational science and the philosophy of science, highlighting how code-level path dependency acquires scientific significance in data-intensive discovery domains.

Mechanisms of Linkage

The theoretical linkage between algorithmic and scientific path dependency in materials AI is realized through four distinct yet interrelated mechanisms that translate early computational choices into enduring scientific constraints.

Table 2 specifies the concrete mechanisms by which early algorithmic choices are translated into downstream scientific constraints, identifying the intervention points at which lock-in can still be disrupted.

Table 2. Mechanisms through which algorithmic path dependency becomes scientific path dependency

Algorithmic trigger	Translating mechanism	Immediate downstream effect in the materials AI workflow	Long-term scientific consequences
Random initialization privileges one optimization trajectory over others	Output filtering	Certain candidates are surfaced, ranked, or discarded early	Some material families underrepresented despite potential
Early training exposure favors particular chemistries or regimes	Output filtering + knowledge representation	Embedding space overrepresents initially encountered patterns	Subsequent hypotheses biased already emphasized regions

Initial hyperparameter settings stabilize one generalization pattern	Resource allocation	Experimental attention follows model outputs judged most promising	Time, system effort, and resource accuracy around path
Descriptor or feature-set choice encodes one view of structure–property relations	Knowledge representation	Some physical relations become legible while others remain invisible	The formal vocabulary narrows or represents a scheme
Early stopping criteria freeze a partially explored solution landscape	Output filtering	Candidate rankings reflect prematurely stabilized model behavior	Discovery trajectories become anchored in shallow optima
Repeated publication of algorithmically favored successes	Publication bias	Positive results from one pipeline become a visible benchmark	Alternative approaches lose legitimate attention to cumulative citation
Experimental validation follows path-dependent rankings	Resource allocation + publication bias	Validated outputs feed back into future datasets and benchmark norms	Scientific forms reflect historical algorithmic contingencies
Community reuse of dominant datasets and pipelines	All four mechanisms jointly	Standardized workflows become the default discovery infrastructure	Epistemic lock-in happens across journals and functional systems

Output Filtering. Algorithmic path dependency determines which material candidates or property predictions are surfaced for human consideration [14]. Because models sensitive to initialization and data ordering prioritize certain

compositional spaces over others [8, 12], many potentially valuable materials remain invisible to researchers. In active learning pipelines for materials discovery, early algorithmic decisions filter the candidate pool, meaning entire families of compounds are never experimentally pursued simply because they were not generated or ranked highly by the initial model configuration [7, 26].

Resource Allocation. Once filtered outputs are available, they directly influence the allocation of scarce experimental resources and funding. Algorithmic recommendations that emerge from path-dependent training processes guide which materials receive synthesis and characterization efforts [7]. This creates a self-reinforcing cycle whereby materials favored by the initial algorithmic path attract disproportionate attention and investment. In contrast, alternative materials are deprioritized, amplifying the scientific consequences of seemingly technical choices [10].

Knowledge Representation. The descriptors, embeddings, and internal representations learned by the algorithm shape what relationships can be discovered and articulated scientifically [5, 13]. When particular feature sets or graph representations are selected early and become locked in, they determine the conceptual vocabulary available for interpreting structure-property relationships. Alternative physical insights that might arise from different representations are structurally foreclosed, constraining the scientific theories that can be developed around AI-generated discoveries [18].

Publication Bias. Algorithmically driven successes are more likely to be published and cited, reinforcing the dominant path while alternative approaches or negative results remain undocumented [19]. This publication dynamic creates feedback loops where path-dependent algorithmic outcomes become the visible scientific record, making it increasingly difficult for researchers to justify exploration of different algorithmic or scientific directions [23].

Conceptually, these mechanisms can be represented as a flow from algorithmic decision nodes (initialization, data ordering, descriptor choice) through the four mechanisms downward to scientific constraint nodes (question selection, material prioritization, epistemic lock-in), with thickening arrows indicating increasing returns and upward feedback loops showing co-evolution between computation and laboratory practice.

Derived Properties and Corollaries

From the central theoretical claim that algorithmic choices in materials AI create path dependency—where early decisions about initialization, data ordering, and descriptor sets irreversibly shape subsequent search trajectories—three important corollaries follow, each revealing structural constraints on scientific discovery. Corollary 1 asserts that the diversity of materials AI research outputs is inherently constrained by algorithmic path dependency, meaning that because early computational choices narrow both the explored parameter space and the forms of representable knowledge, the field collectively produces a less diverse set of discoveries than would be possible under genuine methodological pluralism, effectively biasing results toward known phases or easily accessible local minima [3, 26]. Corollary 2 argues that reversing scientific path dependency requires addressing algorithmic path dependency directly. Yet, this reversal is often practically impossible once lock-in has occurred, as high switching costs in legacy codebases, trained models, normalized descriptor sets, and established research programs create strong economic and social incentives to persist with suboptimal but entrenched methods [1, 2, 8]. Corollary 3 therefore prescribes that methodological pluralism—explicitly employing multiple random initializations, varied data orderings, diverse descriptor sets, and parallel algorithmic approaches such as different optimization strategies or surrogate models—is essential to mitigate path-dependent effects and genuinely expand the scientific possibility space in materials discovery, transforming what might otherwise be a deterministic narrowing into a more exploratory and robust search [11, 28, 29]. Taken together, these corollaries reveal that path dependency is not merely a technical nuisance to be optimized away. Still, rather a structural feature of AI-driven materials research with profound epistemic implications: it determines not only what gets discovered but what can be discovered, shaping the very boundaries of reproducible knowledge.

Implications for Materials AI Practice

This theoretical analysis carries direct implications for research practice. Authors should explicitly report path-dependent choices such as initialization seeds and data ordering, conduct sensitivity analyses across alternative starting conditions, and acknowledge how their algorithmic

decisions constrain the generalizability of findings [3, 15]. Reviewers ought to probe manuscripts for awareness of path dependency, question results based on single initializations, and request robustness checks against different algorithmic paths. At the community level, the field would benefit from developing standardized reporting guidelines for path-dependent decisions, incentivizing methodological pluralism through funding and publication norms, and retrospectively studying historical path dependencies in major materials AI breakthroughs [7, 14].

Conclusion

This paper has articulated that algorithmic path dependency propagates into scientific path dependency in materials AI. Early choices in model initialization, data sequencing, and descriptor selection are not merely technical but shape the very trajectory of scientific discovery. By drawing on foundational concepts from economics and applying them to contemporary materials in AI literature, the analysis reveals the need for greater awareness of how computational decisions become epistemic commitments. Recognizing and actively managing path dependency through pluralism and transparency will be critical if materials AI is to realize its full potential for open-ended discovery rather than constrained, self-reinforcing trajectories.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 05 May 2023 Revised: 01 Jul 2023 Accepted: 02 Aug 2023
Published online: 18 January 2024

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Khan SA, Ponce P, Yu Z, Ponce K. Investigating economic growth and natural resource dependence: An asymmetric approach in developed and developing economies. *Resour Policy*. 2022;77:102672.
- Goldfarb A, Tucker C. Digital economics. *J Econ Lit*. 2019;57(1):3-43.
- Brenner T, Zu Jeddelloh S. Path dependence in an evolving system: A modeling perspective. *Cliometrica*. 2024;18(1):1-36.
- 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).
- Maia MA, Rocha IB, Kerfriden P, van der Meer FP. Physically recurrent neural networks for path-dependent heterogeneous materials: Embedding constitutive models in a data-driven surrogate. *Comput Methods Appl Mech Eng*. 2023;407:115934.
- He X, Chen JS. Thermodynamically consistent machine-learned internal state variable approach for data-driven modeling of path-dependent materials. *Comput Methods Appl Mech Eng*. 2022;402:115348.
- Ramprasad R, Batra R, Paliana G, Mannodi-Kanakithodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *NPJ Comput Mater*. 2017;3(1):54.
- Wang AY, Murdock RJ, Kauwe SK, Oliynyk AO, Gurlo A, Brgoch J, et al. Machine learning for materials scientists: An introductory guide toward best practices. *Chem Mater*. 2020;32(12):4954-65.
- Mozaffar M, Bostanabad R, Chen W, Ehmann K, Cao J, Bessa MA. Deep learning predicts path-dependent plasticity. *Proc Natl Acad Sci U S A*. 2019;116(52):26414-20.
- Schleder GR, Padilha AC, Acosta CM, Costa M, Fazzio A. From DFT to machine learning: Recent approaches to materials science—a review. *J Phys Mater*. 2019;2(3):032001.
- Kusne AG, Yu H, Wu C, Zhang H, Hattrick-Simpers J, DeCost B, et al. On-the-fly closed-loop materials discovery via Bayesian active learning. *Nat Commun*. 2020;11(1):5966.
- Vlassis NN, Sun W. Component-based machine learning paradigm for discovering rate-dependent and pressure-sensitive level-set plasticity models. *J Appl Mech*. 2022;89(2):021003.
- Chiodo M, Müller D, Sienknecht M. Educating AI developers to prevent harmful path dependency in AI resort-to-force decision making. *Aust J Int Aff*. 2024;78(2):210-9.
- Sydow J, Schreyögg G, Koch J. On the theory of organizational path dependence: Clarifications, replies to objections, and extensions. *Acad Manag Rev*. 2020;45(4):717-34.
- Fan Z, Xing B, Cao P. Predicting path-dependent diffusion barrier spectra in vast compositional space of multi-principal element alloys via convolutional neural networks. *Acta Mater*. 2022;237:118159.
- Ali SM, Dick S, Dillon S, Jones ML, Penn J, Staley R. Histories of artificial intelligence: A genealogy of power. *BJHS Themes*. 2023;8:1-8.
- Faillace CA, Grunberg RL, Morin PJ. Historical contingency and the role of post-invasion evolution in alternative community states. *Ecology*. 2022;103(7):e3711.

Debray R, De Luna N, Koskella B. Historical contingency drives compensatory evolution and rare reversal of phage resistance. *Mol Biol Evol.* 2022;39(9):msac182.

Nasteski V. An overview of the supervised machine learning methods. *Horizons B.* 2017;4:51-62.

McConwell AK, Turner DD. Historical contingency: A special issue on epistemic & non-epistemic values in historical sciences. *J Philos Hist.* 2023;17(1):1-8.

Ge W, Tagarielli VL. A computational framework to establish data-driven constitutive models for time-or path-dependent heterogeneous solids. *Sci Rep.* 2021;11(1):15916.

Kalenitchenko D, Peru E, Galand PE. Historical contingency impacts on community assembly and ecosystem function in chemosynthetic marine ecosystems. *Sci Rep.* 2021;11(1):13994.

Riebesell J, Goodall RE, Benner P, Chiang Y, Deng B, Ceder G, et al. A framework to evaluate machine learning crystal stability predictions. *Nat Mach Intell.* 2025;7(6):836-47.

Sedik A, Ilyasu AM, Abd El-Rahiem B, Abdel Samea ME, Abdel-Raheem A, Hammad M, et al. Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections. *Viruses.* 2020;12(7):769.

Maia MA, Rocha IB, Kovačević D, Van der Meer FP. Physically recurrent neural network for rate and path-dependent heterogeneous materials in a finite strain framework. *Mech Mater.* 2024;198:105145.

Pandey D, Niwaria K, Chourasia B. Machine learning algorithms: A review. *Mach Learn.* 2019;6(2):916-22.