Physics-Guided Graph Neural Networks for Interpretable Active-Site Evolution Prediction in Water Oxidation Catalysts

Joel Deaiaz
Department of Computer Science, University of Houston, Houston, TX, USA.
hellojoel@uh.edu

Mikkel Powell
School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
mikkel1981@uc.edu

Abstract

Water oxidation is a cornerstone reaction for renewable energy storage and sustainable hydrogen production, yet the rational design of efficient and durable catalysts remains constrained by an incomplete understanding of how active sites dynamically restructure under operating conditions. Physics-guided graph neural networks offer a transformative framework to bridge this gap, integrating first-principles physical constraints with data-driven learning to predict the structural and electronic evolution of catalytic surfaces. This paper presents a systems-level examination of such a hybrid modeling paradigm, focusing on the architecture, interpretability mechanisms, data infrastructure, and deployment strategies that underpin reliable active-site evolution predictions. We discuss how graph neural networks encode atomistic environments as graph structures whose nodes and edges capture coordination chemistry and electronic interactions, while physics-guided regularizations enforce thermodynamic and spectroscopic consistency. Special attention is devoted to interpretability, enabling researchers to trace which structural motifs and electronic signatures drive activity, including operando spectroscopic fingerprints such as the formation of Zhang-Rice singlet states. The discussion extends to data governance, emphasizing the need for FAIR-compliant data ecosystems that amalgamate high-throughput experiments, density functional theory calculations, and real-time spectroscopy. Robustness against distributional shifts, fairness in catalyst discovery across material families, and the sustainability implications of AI-accelerated catalyst pipelines are addressed as cross-cutting concerns. Finally, policy considerations surrounding open science, intellectual property, and equitable access to AI tools for the global energy transition are explored. The paper argues that a tightly coupled socio-technical infrastructure, where domain-informed model architecture and transparent governance coexist, is essential for translating physics-guided graph neural networks from laboratory proof-of-concept to impactful, field-deployable decision-support systems.

Keywords

physics-guided machine learning, graph neural networks, water oxidation, active-site evolution, interpretability, sustainable catalysis.

1. Introduction

The electrochemical splitting of water into oxygen and hydrogen stands as one of the most promising pathways for storing intermittent renewable energy in chemical bonds, yet the

oxygen evolution reaction (OER) involved remains a kinetic bottleneck that demands highly active, stable, and earth-abundant catalysts. Decades of experimental and theoretical research have demonstrated that the performance of OER catalysts is not governed by static surface properties alone but by the dynamic reconstruction of active sites under anodic potentials, where oxidation states, coordination numbers, and local bonding environments continuously evolve. This phenomenon, often observed through operando X-ray absorption spectroscopy and surface-sensitive probes, reveals that pre-catalytic phases can transform into layered oxyhydroxides, amorphous oxides, or mixed-valence motifs, making the direct rationalization of structure-activity relationships exceptionally challenging [1]. Understanding and predicting such active-site evolution is therefore critical for designing next-generation catalysts, yet doing so requires computational frameworks that can capture complex, non-linear interactions across multiple length and time scales while respecting the fundamental laws of physics.

Machine learning has emerged as a powerful complement to traditional electronic structure methods, offering the ability to learn patterns from large datasets and make rapid predictions. Among the various architectures, graph neural networks (GNNs) are particularly well suited for modeling catalytic materials because they naturally represent atoms as nodes and bonds as edges, enabling the extraction of local chemical environments and long-range order simultaneously [2]. However, purely data-driven GNNs often operate as black-box predictors, offering limited physical insight and struggling to generalize when confronted with reaction conditions or material compositions that lie outside their training distribution. Physics-guided machine learning addresses this limitation by embedding known physical principles—such as conservation laws, thermodynamic stability criteria, or spectroscopic signatures—directly into the model training process, thereby constraining predictions to remain within physically plausible manifolds [3]. This fusion of inductive bias and data-driven flexibility provides a route toward not only accurate but also interpretable models of active-site evolution, where predictions can be traced back to specific atomic rearrangements and electronic structure changes.

The present paper undertakes a systems-level analysis of physics-guided GNNs for interpretable active-site evolution prediction in water oxidation catalysts. Rather than focusing narrowly on algorithmic innovations, we examine the broader socio-technical architecture required to design, validate, and deploy such models effectively. This includes the representation of catalyst surfaces as dynamic graphs, the integration of multi-source experimental and computational data, the governance of data and model artifacts, and the interpretation of predictions in a manner that empowers human decision-makers in catalyst development. We further discuss robustness against data distribution shifts, fairness in exploring diverse chemical spaces, and the sustainability implications of computationally intensive AI workflows. By situating the technical discussion within a wider context of infrastructure, policy, and equitable deployment, we argue that the realization of physics-guided GNNs as reliable tools for catalyst discovery demands more than algorithmic excellence; it necessitates a deliberate reconfiguration of the human and institutional systems that generate, share, and act upon scientific knowledge.

2. System Architecture and Graph Representation of Catalytic Surfaces

The design of a physics-guided GNN for active-site evolution prediction begins with the construction of a graph representation that captures the essential structural and electronic features of a catalyst surface at a given snapshot of an operando experiment or density functional theory (DFT) trajectory. Nodes in such a graph typically encode atom-specific

properties including element type, oxidation state, partial charge, and magnetic moment, while edges encode bond distances, bond orders, and coordination types. To account for dynamic reconstruction, the graph must be allowed to evolve: node and edge attributes change as oxidation states shift and ligands are exchanged, and the topology itself may transform when atoms dissolve, migrate, or reconstruct into new phases. This temporal graph structure thus becomes a spatiotemporal abstraction of the catalyst's working state, onto which graph convolution operations are applied to learn representations that respect local chemical symmetry and global connectivity [2].

Physics-guided constraints can be incorporated at multiple levels of this architecture. In the most direct form, energy-based penalties can be imposed on the model's output to ensure that predicted atomic configurations adhere to known thermodynamic equilibria, such as the Pourbaix stability regions of metal oxides under applied potential and pH. Similarly, bond-valence sum rules can serve as a regularizer, discouraging the prediction of unrealistic oxidation state combinations that would violate local charge neutrality. When operando spectroscopic data are available, the GNN can be encouraged—through a multi-task learning objective—to simultaneously predict not only catalytic activity metrics but also spectral features like X-ray absorption near-edge structure (XANES) or extended X-ray absorption fine structure (EXAFS) signatures, linking the latent representation to experimentally observable quantities [6]. This dual prediction task forces the model to learn a chemically meaningful internal representation, because it must explain both the functional output and the spectroscopic fingerprint, thereby reducing the risk of finding spurious correlations.

The architectural choices also reflect a fundamental trade-off between expressiveness and computational tractability. Full message-passing schemes over large periodic supercells can capture long-range elastic and electronic effects but incur high computational costs, which limit their use in high-throughput screening pipelines. On the other hand, local approximations that restrict message passing to a fixed cutoff radius reduce computational load but may miss collective phenomena such as surface reconstruction waves or long-range strain coupling that are known to influence active-site behavior. A system-level design must therefore strike a balance, perhaps by adopting hierarchical graph architectures where local interactions are modeled with fine-grained attention mechanisms and longer-range effects are captured through coarse-grained supernodes representing clusters or entire facets [7]. Attention mechanisms further enable the model to assign differential importance to neighboring atoms, offering a natural pathway toward interpretability that we discuss in later sections. The deployment of such architectures in practice requires not only careful algorithmic design but also robust software engineering practices that ensure reproducibility across different hardware platforms, a point that connects the technical infrastructure to the broader ecosystem of scientific computing.

3. Data Infrastructure and Governance for Catalyst Evolution Data

Reliable training and validation of physics-guided GNNs depend on an underlying data infrastructure that is both rich in chemical diversity and governed by principles of openness and interoperability. The Materials Project, the Open Catalyst Project, and similar initiatives have demonstrated the value of large-scale, openly accessible repositories of DFT-computed material properties, yet capturing active-site evolution under operando conditions demands data types that go far beyond static bulk or slab calculations [5]. Time-resolved spectroscopy, electrochemical mass spectrometry, and differential electrochemical mass spectrometry produce multidimensional signals that must be carefully aligned with atomic-scale structural

models. Furthermore, experimental data are often heterogeneous in format, subject to systematic biases from instrument calibrations, and collected under widely varying conditions of electrolyte composition, temperature, and mass transport. Building an integrated data pipeline that can fuse these disparate sources into a coherent graph-labeled dataset represents a substantial systems engineering challenge, requiring standardized metadata schemas, automated data extraction workflows, and robust provenance tracking.

Governance frameworks such as the FAIR Guiding Principles—making data Findable, Accessible, Interoperable, and Reusable—become essential when multiple research groups and institutions contribute data to a shared model development effort [10]. In the context of catalyst evolution, this means that every experimental run or simulation trajectory should be accompanied by rich metadata describing the synthesis conditions, electrochemical protocol, spectroscopic acquisition parameters, and any post-processing steps. Without such meticulous documentation, models trained on aggregated data risk learning artifacts that stem from inconsistent preprocessing rather than genuine physical phenomena. The governance challenge extends to versioning and quality control: as new experimental techniques emerge and older datasets are re-analyzed, the data infrastructure must support incremental updates while maintaining backward compatibility for model retraining. This is akin to the continuous integration and deployment pipelines used in large-scale software engineering, adapted to the peculiarities of scientific data streams [14]. The socio-technical dimension here is pronounced, because the adoption of such infrastructure requires aligning incentives across academic laboratories, which traditionally prioritize rapid publication over careful data curation, and funding agencies, which are increasingly mandating data management plans.

In addition to experimental data, synthetic data generated by DFT or molecular dynamics play a crucial role in augmenting the limited experimental coverage. Physics-guided GNNs can leverage these simulations to pre-train representations on idealized surfaces, which are then fine-tuned on sparse experimental observations. However, the fidelity of simulated data depends on the choice of exchange-correlation functional, pseudopotential, and solvation model, introducing epistemic uncertainties that must be propagated through the machine learning pipeline. Transparently communicating these uncertainties to downstream users—whether theorists designing new catalysts or experimentalists prioritizing synthesis targets—is an ethical obligation and a practical necessity. Data governance policies must therefore address not only the availability of data but also the clear qualification of its uncertainty and limitations, enabling informed decision-making across the catalyst discovery ecosystem.

4. Interpretability Mechanisms in Active-Site Prediction

A central promise of physics-guided GNNs is their potential to make the active-site evolution process interpretable, transforming the model from a black-box oracle into a tool for scientific insight. Interpretability in this context operates at multiple levels: global interpretability reveals which types of structural motifs or electronic configurations are generally predictive of high activity or stability; local interpretability explains the prediction for a specific catalyst surface at a given time step, identifying the critical atoms, bonds, or electronic states that drive the model's output. Graph attention mechanisms naturally facilitate local interpretability by computing attention coefficients over edges, which can be visualized as heatmaps superimposed on the atomic structure to highlight regions where the model focuses its reasoning [7]. When such attention patterns are correlated with known descriptors—such as the eg orbital filling of surface metal ions or the metal-oxygen covalency in transition metal

oxides—researchers can validate whether the model has learned physically meaningful relationships or has latched onto spurious correlations.

More advanced interpretability techniques, such as perturbation-based explanation methods originally developed for graph neural networks, can be adapted to the catalysis domain. For example, systematically masking individual atoms or functional groups and measuring the change in predicted activity provides a counterfactual analysis akin to experimental site-blocking studies [13]. When such in silico perturbation reveals that the removal of a specific bridging oxygen atom drastically alters the predicted overpotential, it mirrors the experimental observation that certain lattice oxygen sites participate directly in the OER mechanism via lattice oxygen oxidation pathways. The physics-guided nature of the model adds a layer of reliability to these interpretations because the predictions are constrained by spectroscopic and thermodynamic consistency checks, meaning that a highlighted active site must not only correlate with high activity but also be consistent with the oxidation state and coordination changes observed experimentally. This convergence of data-driven attention and physics-based plausibility creates a powerful reasoning loop, enabling researchers to generate hypotheses about active-site dynamics that can be tested through targeted operando experiments.

A particularly compelling illustration arises in systems where Zhang-Rice singlet states form through a two-step oxidation process, triggering exceptionally high water oxidation activity under operando conditions [12]. In such catalysts, the interplay between metal 3d and oxygen 2p orbitals leads to localized hole states that profoundly influence the electronic structure and surface reactivity. A physics-guided GNN, if trained on data that include both structural coordinates and spectroscopic signatures of such singlet states, can learn to associate the emergence of this electronic motif with a sudden increase in catalytic turnover. The interpretability tools can then trace this association back to the specific combination of geometric distortion—such as Jahn-Teller elongation—and ligand coordination that gives rise to the singlet. This capability transforms the model into a hypothesis-generating instrument that augments human intuition, rather than a mere predictive engine. Ensuring that such interpretations are robust requires that the interpretability methods themselves be subjected to rigorous validation, for instance by checking that the explanations remain stable under small perturbations of the input and that they transfer across chemically similar systems. This meta-level of scrutiny reflects a broader research quality infrastructure that is necessary as AI-driven insights increasingly guide experimental investments.

5. Robustness, Uncertainty, and Fairness in Predictive Modeling

Deploying physics-guided GNNs in realistic catalyst discovery workflows demands attention to robustness against the distributional shifts that are endemic to materials science. A model trained on a dataset dominated by iridium and ruthenium oxides may fail silently when applied to first-row transition metal oxides, not because the underlying physics differs fundamentally, but because the latent space of the learned representations is poorly calibrated outside the training manifold. Physics-guided regularization can mitigate this to some extent by anchoring predictions to universal principles, but it cannot eliminate epistemic uncertainty entirely. Quantifying predictive uncertainty—through ensemble methods, Monte Carlo dropout, or Bayesian graph neural networks—is therefore an essential component of a trustworthy system [18]. When the model expresses high uncertainty for a particular catalyst composition or reaction condition, this signal can be used to trigger an active learning loop, prompting the execution of additional DFT calculations or targeted operando experiments to

enrich the dataset in that underrepresented regime. This tight coupling between model uncertainty and experimental design exemplifies the closed-loop automation that accelerates materials discovery while maintaining scientific rigor.

Fairness in AI-assisted catalyst discovery is a less frequently discussed but equally important consideration. The exploration of chemical space for water oxidation catalysts has historically been biased toward precious metals and a narrow set of oxide families, partly due to the legacy of experimental ease and the path dependence of research funding. When machine learning models are trained on such historically biased datasets, they risk perpetuating and even amplifying the neglect of alternative material classes such as polyoxometalates, chalcogenides, or high-entropy oxides that might ultimately yield more sustainable and cost-effective catalysts [15]. Physics-guided architectures offer a unique corrective opportunity because they do not rely solely on data abundance; they can incorporate domain knowledge about viable oxidation states, coordination preferences, and thermodynamic constraints that extend across the entire periodic table. By encoding this knowledge as inductive bias, the model can make physically informed extrapolations into underrepresented regions of chemical space, thereby broadening the scope of exploration in a more equitable manner. However, this potential can only be realized if the data governance frameworks discussed earlier actively prioritize data collection from diverse chemical systems and if the evaluation metrics for model performance explicitly account for predictive equity across material families, rather than merely optimizing aggregate accuracy.

The fairness lens also extends to the human stakeholders involved. Predictive models that are developed solely within well-resourced institutions may not address the catalytic challenges most relevant to water oxidation in the context of decentralized energy systems in the Global South, where access to purified water, specific electrolytes, or noble-metal catalyst precursors may be limited. Engaging a globally diverse community of researchers in defining the target material space and performance criteria is a governance strategy that aligns the AI development process with broader sustainability and equity goals. Such participatory approaches, while adding complexity to project management, ultimately build the socio-technical resilience required for widespread technology adoption.

6. Deployment, Scalability, and Sustainability in Clean Energy Applications

The translation of a physics-guided GNN from a research prototype to a field-deployable tool for catalyst development involves a series of infrastructural and operational decisions that determine its scalability and real-world impact. A primary consideration is whether the model should be deployed as a cloud-based service accessible via web interfaces and application programming interfaces, or as a lightweight edge model that can run on local computing clusters within experimental facilities. Cloud deployment offers advantages in terms of centralized model updates, version control, and the ability to leverage elastic computing resources for intensive inference or retraining jobs. However, it raises concerns about data privacy when proprietary catalyst formulations are evaluated, and about connectivity dependencies in regions with unreliable internet access. Edge deployment, on the other hand, requires compressing and optimizing the GNN to run efficiently on modest hardware, which may compromise predictive accuracy or the granularity of interpretability outputs. A hybrid architecture, where lightweight models serve as rapid screening filters and detailed physics-guided analyses are performed in the cloud on demand, represents a balanced design that can accommodate diverse user profiles, from academic groups to industrial R&D laboratories [14].

Sustainability of the AI lifecycle itself must also be scrutinized. Training large graph neural networks on extensive catalyst datasets can incur significant energy consumption and carbon emissions, potentially undercutting the very environmental benefits that improved water oxidation catalysts aim to deliver [16]. System designers must therefore weigh the computational cost of model development against the catalytic efficiency gains it enables. Strategies such as transfer learning, where a backbone GNN pre-trained on a broad materials dataset is fine-tuned for specific catalyst families, can dramatically reduce the required training energy. Moreover, the use of physics-informed loss functions can accelerate convergence and reduce the need for exhaustive hyperparameter searches, because the model is guided toward physically plausible regions of parameter space from the outset. Reporting the full lifecycle environmental impact of AI-driven catalyst discovery should become a standard practice, akin to the reporting of turnover frequencies and Faradaic efficiencies for catalysts themselves, to ensure that the digital infrastructure supporting the energy transition does not become an unacknowledged contributor to the problem it seeks to solve.

The scalability of the approach also hinges on the automation of experiment-model feedback loops. High-throughput electrochemical testing platforms, when integrated with automated X-ray absorption spectroscopy beamlines, can generate time-resolved data streams that directly feed into the GNN's continuous learning pipeline. This integration requires robust middleware that translates raw detector signals into graph-formatted inputs in real time, applying calibration corrections and extracting spectral features on the fly. The engineering of such automated workflows demands interdisciplinary collaboration between beamline scientists, software architects, and machine learning engineers, extending the system's boundaries well beyond the model itself. As national laboratories and synchrotron facilities increasingly offer remote access and automated experimentation capabilities, the prospect of fully autonomous "self-driving laboratories" for catalyst evolution becomes tangible. The governance of these autonomous systems, including fail-safe mechanisms when the model encounters unfamiliar chemical regimes, becomes a critical area of research that straddles safety engineering and AI ethics.

7. Policy and Ethical Implications of AI-Accelerated Catalyst Discovery

The accelerating integration of physics-guided GNNs into catalyst research generates policy challenges that intersect with intellectual property law, open science norms, and the global governance of clean energy technologies. On one hand, the predictive capabilities of these models could dramatically shorten the innovation cycle for new catalysts, potentially enabling rapid deployment of electrolyzers that reduce the cost of green hydrogen and decarbonize hard-to-abate industrial sectors. On the other hand, the concentration of model development expertise and proprietary datasets within a few technology companies or leading research institutions risks creating a digital divide in catalytic science. If access to the most powerful physics-guided GNNs is gated by licensing fees or trade secrets, smaller research groups and those in lower-income countries may be excluded from the frontier of catalyst design, thereby reinforcing existing inequities in the global energy transition. Policy frameworks that incentivize the open release of pre-trained models and benchmark datasets, while protecting legitimate commercial interests through time-limited exclusive use periods or patent pools, could balance the competing demands of innovation and access [19, 20].

The interpretability of physics-guided GNNs also bears on regulatory and safety governance. When a GNN-guided catalyst design is selected for scale-up and integration into a commercial electrolyzer, regulatory bodies will require assurance that the active-site

evolution predictions are reliable, that failure modes are understood, and that the risk of hazardous degradation products is minimal. Interpretable models that can provide mechanistic explanations for their predictions, coupled with well-quantified uncertainty estimates, are far more likely to satisfy such regulatory scrutiny than opaque deep learning architectures. This creates a policy incentive to mandate interpretability not merely as a scientific aspiration but as a component of the safety case for AI-designed materials. The development of standardized protocols for validating the interpretability claims of machine learning models in chemistry—analogous to the Good Machine Learning Practice guidelines emerging in medical device regulation—may become necessary [23]. Multi-stakeholder initiatives that bring together catalyst manufacturers, academic researchers, regulatory agencies, and civil society representatives can co-design such protocols, ensuring they are both scientifically rigorous and socially legitimate.

Finally, the workforce implications of AI-accelerated catalyst discovery deserve attention. As machine learning systems take on roles traditionally performed by human researchers—sifting through literature, proposing candidate structures, and even designing experiments—the skill profiles required in catalysis research will shift. Educational programs will need to evolve to equip the next generation of scientists with fluency in both physical chemistry and machine learning systems design, as well as the ethical reasoning skills to navigate the dual-use potential of powerful predictive models. Funding agencies and universities have a responsibility to support this transition through interdisciplinary training grants, curriculum redesign, and the creation of career paths that value data stewardship and software engineering alongside traditional academic metrics. The long-term success of physics-guided GNNs in water oxidation catalysis thus rests not only on algorithmic breakthroughs but on a deliberate and inclusive reimagination of the human infrastructure that supports scientific discovery.

8. Conclusion

Physics-guided graph neural networks represent a paradigmatic convergence of domain knowledge and data-driven modeling, uniquely positioned to illuminate the complex, dynamic evolution of active sites in water oxidation catalysts. This paper has argued that realizing the full potential of such models requires a systems-level perspective that extends far beyond the neural architecture itself. It demands the construction of interoperable data infrastructures that faithfully capture the richness of operando experiments and first-principles simulations, governed by FAIR principles to ensure longevity and reusability. It requires the deliberate embedding of interpretability mechanisms that empower human researchers to trace predictions back to physically meaningful atomic and electronic descriptors, thereby building trust and enabling mechanistic discovery. It further demands attention to robustness under distributional shift, fairness in chemical space exploration, and sustainability in the computational lifecycle of AI models.

The deployment of these models will occur within a complex socio-technical landscape where policy, intellectual property, regulatory safety, and global equity intersect. Ensuring that physics-guided GNNs contribute to a just and rapid energy transition calls for proactive governance that promotes open innovation while safeguarding against the concentration of AI capabilities. The interpretability of these models is not merely a technical convenience but a governance necessity, providing the transparency required for regulatory approval and public trust. As the field moves toward self-driving laboratories that couple high-throughput experimentation with real-time machine learning inference, the need for resilient, ethically

grounded, and collaboratively designed AI systems becomes paramount. Ultimately, the physics-guided graph neural network is not just a computational tool; it is a node in a larger infrastructure of knowledge production, whose design must reflect the values and priorities of the global community it seeks to serve.

References

1. Seh, Z. W., Kibsgaard, J., Dickens, C. F., Chorkendorff, I., Nørskov, J. K., & Jaramillo, T. F. (2017). Combining theory and experiment in electrocatalysis: Insights into materials design. Science, 355(6321), eaad4998.

2. Xie, T., & Grossman, J. C. (2018). Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. Physical Review Letters, 120(14), 145301.

3. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707.

4. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

5. Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., ... & Persson, K. A. (2013). Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. APL Materials, 1(1), 011002.

6. Bergmann, A., & Roldan Cuenya, B. (2019). Operando insights into nanoparticle transformations during catalysis. ACS Catalysis, 9(11), 10020–10043.

7. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2018). Graph attention networks. International Conference on Learning Representations.

8. Seitz, L. C., Dickens, C. F., Nishio, K., Hikita, Y., Montoya, J., Doyle, A., ... & Jaramillo, T. F. (2016). A highly active and stable IrOx/SrIrO3 catalyst for the oxygen evolution reaction. Science, 353(6303), 1011–1014.

9. Nørskov, J. K., Bligaard, T., Rossmeisl, J., & Christensen, C. H. (2009). Towards the computational design of solid catalysts. Nature Chemistry, 1(1), 37–46.

10. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(1), 1–9.

11. Kitchin, J. R. (2018). Machine learning in catalysis. Nature Catalysis, 1(4), 230–232.

12. Peng, C. K., Lin, Y. C., Chiang, C. L., Qian, Z., Huang, Y. C., Dong, C. L., ... & Lin, Y. G. (2023). Zhang-Rice singlets state formed by two-step oxidation for triggering water oxidation under operando conditions. Nature Communications, 14(1), 529.

13. Ying, R., Bourgeois, D., You, J., Zitnik, M., & Leskovec, J. (2019). Gnnexplainer: Generating explanations for graph neural networks. Advances in Neural Information Processing Systems, 32.

14. Baylor, D., Breck, E., Cheng, H. T., Fiedel, N., Foo, C. Y., Haque, Z., ... & Page, D. (2017). TFX: A TensorFlow-based production-scale machine learning platform.

Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1387–1395.

15. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1–35.
16. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. arXiv preprint arXiv:1906.02243.
17. Corma, A., & Serna, P. (2010). High-throughput techniques for the discovery of new catalytic materials. Chemical Society Reviews, 39(5), 1541–1556.
18. Tran, K., Neiswanger, W., Yoon, J., Zhang, Q., Xing, E., & Ulissi, Z. W. (2020). Methods for comparing uncertainty quantifications for materials property predictions. Machine Learning: Science and Technology, 1(2), 025006.
19. International Energy Agency. (2017). Digitalization and Energy. IEA Publications.
20. Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1).
21. Chanussot, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., ... & Ulissi, Z. (2021). Open Catalyst 2020 (OC20) dataset and community challenges. ACS Catalysis, 11(10), 6059–6072.
22. Jiménez-Luna, J., Grisoni, F., & Schneider, G. (2020). Drug discovery with explainable artificial intelligence. Nature Machine Intelligence, 2(10), 573–584.
23. Fabbri, E., Habereder, A., Waltar, K., Kötz, R., & Schmidt, T. J. (2014). Developments and perspectives of oxide-based catalysts for the oxygen evolution reaction. Catalysis Science & Technology, 4(11), 3800–3821.