ARTIFICIAL INTELLIGENCE IN CHEMISTRY, MATERIALS, AND CATALYSIS: METHODS, MILESTONES, AND OPEN CHALLENGES

Niraj Kumar Sharma
Assistant Professor of Chemistry, Vallabh Govt College Mandi (H.P)
nirajksharma0170@gmail.com

Abstract
Artificial intelligence (AI) has rapidly moved from a promising add-on to a central driver of discovery in chemistry, materials science, and catalysis. Data-centric workflows, graph and equivariant neural networks, and increasingly capable language-and-tool agents now accelerate prediction, design, and even autonomous experimentation. This paper surveys core representations and model classes; benchmarks and datasets that anchor progress; landmark results including large-scale inorganic materials discovery and self-driving laboratories; and domain-specific advances in reaction prediction, retrosynthesis, and heterogeneous catalysis. We close with limitations—data fidelity, uncertainty, synthesis bottlenecks, and reproducibility—and outline practical steps toward trustworthy, closed-loop discovery.

1. Introduction

Modern chemical and materials discovery faces combinatorial spaces (10^60 molecules, millions of possible inorganic crystals, and vast catalyst–adsorbate–condition combinations). AI narrows these spaces by learning structure–property relationships, guiding inverse design, and orchestrating automated experiments. In 2023–2024, two milestones crystallized the promise and the caveats: Google DeepMind's GNoME scaled deep learning to propose ~2.2 million hypothetical crystals and ~381k predicted-stable compositions, while Lawrence Berkeley National Laboratory's A-Lab converted algorithmic suggestions into 41 successful syntheses out of 58 candidates in 17 days—demonstrating closed-loop acceleration but also exposing synthesis and validation limits.

2. Molecular and Materials Representations

Strings and graphs. Early neural models encoded molecules as SMILES strings; robustness issues motivated SELFIES, a 100% valid molecular string language that improved generative design by ensuring any produced token sequence maps to a valid molecule.

Crystal graphs. For inorganic solids, Crystal Graph Convolutional Neural Networks (CGCNN) represent periodic structures as atom-bond graphs with periodic edges, enabling supervised learning of DFT-level properties. CGCNN and successors underlie many state-of-the-art predictors and appear widely in workflows and codebases.

Equivariance and physics priors. For atomistic energies, forces, and adsorption geometries, E(3)-equivariant architectures (e.g., SchNet variants, GemNet-OC, eSCN) enforce rotational and translational symmetries, dramatically improving data efficiency in materials and catalysis tasks (see §5 datasets OC20/OC22).

3. Model Classes and Tasks

Property prediction. Supervised GNNs and message-passing models predict band gaps, formation energies, elastic constants, and adsorption energies with accuracy competitive with or complementing DFT, especially when trained on large curated sets. Matbench standardized 13 inorganic tasks to compare algorithms and reduce benchmark leakage. Generative and inverse design. Diffusion/VAEs/transformers over SELFIES or graphs enable constrained generation (e.g., target property-guided molecules or crystals). For surfaces, task-specific GNNs estimate adsorption energies orders of magnitude faster than ab initio, enabling broader screens.

Language-and-tool agents. LLM agents augmented with cheminformatics, simulation, and literature tools (e.g., ChemCrow) can plan syntheses, design experiments, and execute multi-step reasoning, while newer evaluations probe chemical knowledge and reasoning limits. Autonomous labs. Active-learning loops couple predictors with robotics and characterization (e.g., A-Lab), turning AI suggestions into validated compounds and feeding results back into models.

4. AI for (Small-Molecule) Chemistry

Forward reaction prediction. The Molecular Transformer framed product prediction as sequence-to-sequence translation with calibrated uncertainty, achieving state-of-the-art accuracy and practical confidence estimates. This catalyzed a wave of transformer-based models for yield, selectivity, and condition prediction. Retrosynthesis and route planning. Template-based, template-free, and hybrid methods continue to advance; recent work targets interpretability (e.g., RetroExplainer) and self-improving search (e.g., ReSynZ), while reviews consolidate best practices and pitfalls (data bias, action spaces, evaluation). LLMs as lab copilots. ChemCrow integrates 10+ expert tools

with an LLM to carry out tasks spanning synthesis planning to materials design; related systems (e.g., Coscientist) show autonomous experimental design and execution with tool use, but reliability and safeguards remain active concerns.

5. AI for Materials Discovery

Benchmarks and leaderboards. Matbench (and Matbench-Discovery) provide standardized tasks and train/test splits for fair comparison, analogous to ImageNet's role in CV, and have become the default yardstick for solid-state prediction models. Scaling laws and foundation models for crystals. GNoME demonstrated that scaling data and model capacity can expand the known thermodynamic convex hull by an order of magnitude. However, predictions alone are not sufficient: stability ≠ synthesizability, and reproducibility depends on experimentalcontext. The A-Lab's high conversion rate offers a template for closing this gap via active learning and robotics. Independent news and perspectives highlight both excitement and over-claim risks.

Open data and community challenges. Open Catalyst Project (OCP) organizes large-scale community benchmarks (OC20/OC22) for surface–adsorbate energetics with S2EF and IS2RE tasks, spurring progress in equivariant GNNs and data-efficient training; related leaderboards and challenges continue to raise the bar.

6. AI in (Heterogeneous) Catalysis

From descriptors to learned energetics. Traditional volcano-plot descriptors (e.g., adsorption energies) can be learned directly with GNNs, enabling rapid exploration of catalyst composition, structure, and adsorbate coverage without exhaustive DFT. GAME-Net and related models deliver fast adsorption-energy estimates for large organics on surfaces. Comprehensive reviews and perspectives. Recent reviews summarize how ML helps bridge the "complexity gap" between idealized calculations and realistic catalytic interfaces (defects, solvents, potentials), document ML-accelerated simulations (ML potentials, enhanced sampling), and assess HTE+ML strategies. Equivariant models & datasets. OCP's OC20/OC22 datasets define adsorption-energy and force-prediction tasks across millions of DFT calculations, establishing shared metrics and training regimes (e.g., joint training improvements, learning long-range interactions in oxides). Emerging frontiers. Generative AI for catalyst design, LLM-assisted experimentation, and domain-specific tools for single-atom catalysts are expanding the design space; however, validation and mechanistic interpretability remain gating factors for practical deployment.

7. Reliability, Reproducibility, and Responsible Use

Data quality and shift. Reaction datasets (text-mined, proprietary) and materials repositories (computed vs. experimental) contain biases and inconsistencies; models can silently exploit artifacts. Community benchmarks (Matbench, OCP) and stronger dataset documentation help, but domain shift to real-world conditions (impurities, morphology, electrolytes) is substantial. Uncertainty and calibration. Calibrated probabilities (e.g., Molecular Transformer's uncertainty estimator) and conformal prediction can triage risky model recommendations and guide active learning. Synthesis bottlenecks. Large-scale proposals (e.g., GNoME) must pass through the sieve of synthesizability and utility; autonomous labs reduce cycle time but still face reagent handling, phase purity, and characterization challenges. Balanced media coverage and post-publication critique caution against hype. Human-in-the-loop. The most robust systems pair models with expert priors, mechanistic reasoning, and interpretable features (e.g., mechanistically grounded reaction descriptors), with LLM agents serving as copilots rather than replacements.

8. Practical Guidance for Teams

1. Start with the right representation and benchmark. For molecules, consider SELFIES for generation and graph encodings for prediction; for crystals, use CGCNN-style graphs or equivariant models. Validate on Matbench-style splits; for catalysis, adopt OC20/OC22 task definitions.
2. Exploit active learning. Prioritize experiments that maximally reduce uncertainty; close the loop with robotics where possible (or virtual HTE).
3. Use agents carefully. LLM-and-tools systems (e.g., ChemCrow) can draft routes and automate routine tasks; enforce guardrails, audit trails, and human review.
4. Report uncertainty and provenance. Track dataset origin, preprocessing, and confidence; prefer physically consistent models (equivariance, constraints).

5. Plan for synthesis and scale-up. Couple model-suggested targets with retrosynthetic feasibility, substrate/cost constraints, and process considerations; test with small standardized campaigns before scaling.

9. Outlook

The trajectory is clear: richer multimodal data (text, spectra, microscopy), larger and more physically grounded models, and tighter automation will continue to compress the idea-to-material timescale. Yet progress depends on rigorous benchmarks, honest uncertainty, and mechanisms to turn predictions into validated function. The most impactful systems will not simply "find" new molecules or materials—they will explain why they work, make them reliably, and measure them reproducibly.

References:

1. Xie, T. & Grossman, J. C. "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties," Phys. Rev. Lett. 120, 145301 (2018).
2. Xie, T. (CGCNN repository). "Crystal graph convolutional neural networks." GitHub.
3. Dunn, A. et al. "Benchmarking materials property prediction with Matbench and Automatminer," npj Comput. Mater. 6, 138 (2020).
4. Matbench website and leaderboard.
5. Krenn, M. et al. "SELFIES: A 100% robust molecular string representation," arXiv:1905.13741 (2019); perspectives on SELFIES.
6. Schwaller, P. et al. "Molecular Transformer: Uncertainty-Calibrated Chemical Reaction Prediction," ACS Cent. Sci. 5, 1572–1583 (2019).
7. Zhong, Z. et al. "Recent advances in deep learning for retrosynthesis," WIREs Comput. Mol. Sci. (2024).
8. Lee, A. et al. "Retrosynthesis Zero (ReSynZ): self-improving global synthesis planning," J. Chem. Theory Comput. (2024).
9. Bran, A. M. et al. "ChemCrow: Augmenting LLMs with chemistry tools," Nat. Mach. Intell. 6, 1023–1034 (2024).
10. Park, C. W. et al. "The Open Catalyst 2022 (OC22) Dataset and Challenges for Oxide Electrocatalysts," arXiv:2206.08917 (2022).
11. Open Catalyst Project: overview, challenge, and leaderboards.
12. Noh, J. et al. "Fast evaluation of adsorption energy of organic molecules on metal surfaces with GAME-Net," Nat. Comput. Sci. 3, 164–172 (2023).
13. Gdeisat, M. et al. "Bridging the complexity gap in computational heterogeneous catalysis with machine learning," Nat. Catal. 6, 534–548 (2023).
14. Wei, H. et al. "Machine-learning atomic simulation for heterogeneous catalysis," npj Comput. Mater. 8, 230 (2022).
15. Khan, S. et al. "Artificial Intelligence, Machine Learning and High-Throughput Experimentation for Catalysis," ACS Catal. (2024).
16. Merchant, A. et al. "Scaling deep learning for materials discovery," Nature 624, 80–86 (2023). (GNoME)
17. DeepMind blog: "Millions of new materials discovered with deep learning." (GNoME explainer).
18. Ament, S. et al. "An autonomous laboratory for the accelerated synthesis of novel materials," Nature 624, 94–101 (2023). (A-Lab)
19. Chemistry World. "New analysis raises doubts over autonomous lab's materials discoveries." (context and critique).
20. Nature Reviews Chemistry (collection & reviews) on ML for heterogeneous catalysis and HT data.
21. Bozal-Ginesta, C., Pablo-García, S., Aspuru-Guzik, A. "Developing machine learning for heterogeneous catalysis with experimental and computational data." (Review listing).
22. Jablonka, K. M., Krishnan, N. M. A. "On incentives and benchmarks for meaningful ML in chemistry." Nat. Rev. Chem. (2025 Comment).
23. Nature (News/Focus): "Machine Learning in Catalysis" collection.
24. Yu, T., Boob, A. G., Zhao, H. "Machine learning-enabled retrobiosynthesis of molecules," Nat. Catal. 6, 2023 (context on biosynthetic pathways).