

AI-Driven and Quantum-Informed Design of Functional Nanomaterials: Bridging Catalysis, Drug Discovery, and Sustainable Environmental Remediation

Sadiq Khan¹, Faizan Ali², Waheed Zaman Khan^{3,4*}, Muhammad Adnan⁵, Raja Muhammad Jawad Naveed⁶, Badar Rasool⁷, Ulfat Ayub⁸, Muhammad Adeel⁹, Muhammad Raza Malik¹⁰,

¹Department of Chemistry, University of Malakand, Pakistan

²Department of Mathematics, COMSATS University Islamabad (Attock Campus), Attock, Pakistan

³Phystech School of Aerospace Technology, Moscow Institute of Physics and Technology (MIPT), 9 Institutskiy per., Dolgoprudny 141701, Russian Federation

⁴Department of Physics, Division of Science & Technology, University of Education, Lahore 54770, Pakistan

⁵Department of Chemistry, Division of Science & Technology, University of Education, Lahore 54770, Pakistan

⁶Department of Chemistry, University of Gujrat, Punjab 50700, Pakistan

⁷Department of Wildlife and Ecology, University of Veterinary & Animal Sciences (UVAS), Lahore 54000, Pakistan

⁸Department of Fisheries & Aquaculture, University of Veterinary & Animal Sciences (UVAS), Lahore 54000, Pakistan

⁹Department of Physics, University of Sargodha, Punjab, Pakistan

¹⁰School of Chemistry, University of the Punjab, Lahore, Punjab, Pakistan

DOI: <https://doi.org/10.36347/sjpm.2025.v12i09.002>

| Received: 13.09.2025 | Accepted: 01.11.2025 | Published: 22.11.2025

*Corresponding author: Waheed Zaman Khan

Department of Physics, Division of Science and Technology, University of Education, Lahore 54770, Pakistan

Abstract

Review Article

The convergence of machine intelligence, quantum-accurate simulation, and laboratory automation is reshaping how functional nanomaterials are conceived, validated, and deployed across chemistry, medicine, and environmental engineering. This review synthesizes an end-to-end “data-to-device” framework for the AI-driven and quantum-informed design of nanomaterials that bridges three application pillars: (i) green catalysis for clean energy and circular chemistry, (ii) drug discovery and nano-enabled therapeutics, and (iii) sustainable environmental remediation. We survey inverse-design workflows that combine generative models, uncertainty-aware predictors, Bayesian optimization, and active learning with electronic-structure engines (DFT, GW/BSE), free-energy methods (FEP/TI), and machine-learned interatomic potentials to span accuracy-throughput trade-offs via multi-fidelity strategies. On the materials side, we map tunable design spaces single-atom catalysts, 2D/defect-engineered surfaces, porous frameworks (MOFs/COFs), quantum dots, membranes, and bio-hybrids linking structure, defects, and interfacial physics to catalytic turnover, molecular recognition, transport, and durability. For catalysis, we outline pipelines that couple adsorption-energy maps and microkinetics to target CO₂ reduction, OER/ORR, and selective oxidations; for therapeutics, we integrate target modeling, generative ideation, physics-based ΔG estimation, and ADMET triage with synthesis-aware constraints; for remediation, we align pollutant fingerprints with adsorption, photocatalysis, electrocatalysis, and membrane routes while tracking leaching and secondary byproducts. Throughout, we emphasize rigorous reporting reproducible data splits, calibrated uncertainty, and minimum information for models and experiments together with life-cycle assessment, techno-economic analysis, and green-chemistry metrics (e.g., PMI, E-factor) to ensure net-positive impact. We close with a roadmap for closed-loop, self-driving laboratories; interoperable data/metadata standards; and prize-style community benchmarks aimed at delivering trustworthy, scalable, and sustainable nanomaterials from computational blueprints to field and clinical realities.

Keywords: AI-driven materials design, Quantum-informed simulation, Functional nanomaterials, Green catalysis Drug discovery, Environmental remediation.

Copyright © 2025 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

Citation: Sadiq Khan, Faizan Ali, Waheed Zaman Khan, Muhammad Adnan, Raja Muhammad Jawad Naveed, Badar Rasool, Ulfat Ayub, Muhammad Adeel, Muhammad Raza Malik. AI-Driven and Quantum-Informed Design of Functional Nanomaterials: Bridging Catalysis, Drug Discovery, and Sustainable Environmental Remediation. Sch J Phys Math Stat, 2025 Nov 12(9): 389-418.

1. INTRODUCTION

The discovery of functional nanomaterials is undergoing a phase shift powered by three converging forces: scalable artificial intelligence (AI), increasingly practical quantum simulation, and laboratory automation that closes the loop between hypothesis, synthesis, and measurement. In combination, these forces promise not only faster discovery but discovery that is explicitly aligned with sustainability optimizing for energy use, waste, and eco-toxicity from the very first design iterations rather than auditing impacts at the end. “Self-driving” laboratories now integrate autonomous experiment planning with robotic synthesis and in-situ characterization, compressing iteration cycles from months to days and demonstrating order-of-magnitude gains in throughput and optimization efficiency (Tom *et al.*, 2024). At the same time, foundation models trained on chemistry and materials corpora are generalizing across tasks—from property prediction and reaction planning to multi-step synthesis suggestions—bringing zero/low-shot capabilities into lab workflows (Pyzer-Knapp *et al.*, 2025). In parallel, quantum algorithms are maturing to address strongly correlated electrons and excited-state chemistry, two long-standing bottlenecks for catalytic and therapeutic design (Paudel *et al.*, 2022; Weidman *et al.*, 2024). Our central thesis is that the intersection of these developments enables a closed-loop, sustainability-aware pipeline for discovering nanomaterials in catalysis, therapeutics, and environmental remediation—evaluated in hard units of energy, speed, and minimized hazard rather than in isolated accuracy metrics (Nizam *et al.*, 2021).

A unified view across chemistry, physics, and the environment is timely for two reasons. First, AI at scale is delivering qualitatively new capabilities. Recent perspectives document rapid gains in transferability and task coverage for foundation models, including in-context learning and agentic planning for materials tasks that previously required bespoke training (Pyzer-Knapp *et al.*, 2025). These models can constrain candidate spaces with domain-aware priors, propose informative experiments, and calibrate uncertainty so that each robotic run maximizes expected information gain. Second, automation has moved beyond concept demos to robust, reconfigurable platforms. End-to-end loops—where algorithms propose experiments, robots execute them, and measured outcomes update the models—are accelerating the scientific method itself and improving reproducibility by standardizing procedures and metadata capture (Tom *et al.*, 2024).

These breakthroughs coincide with intensifying imperatives for greener development. Life-cycle-assessment (LCA) appraisals of nanomaterials have highlighted inconsistent functional units, incomplete background datasets, and heterogeneous toxicity characterization, creating comparability gaps across studies (Nizam *et al.*, 2021). Nevertheless, consensus is forming around adopting standardized characterization

models such as USEtox®, declaring system boundaries early, and reporting uncertainty so that design decisions can be optimized, not merely audited after the fact. Embedding such sustainability metrics directly into AI and automation utilities—e.g., penalizing solvent and precursor hazards, minimizing process energy per informative measurement—aligns optimization with environmental objectives from the outset (Nizam *et al.*, 2021).

The scope of this review spans four communities that increasingly share one discovery pipeline. For materials scientists and chemists, we synthesize how AI and quantum tools can be made practically useful when coupled to automation and standardized environmental metrics. For pharmacologists, we emphasize how generative and predictive models for nanoscale delivery systems (e.g., carriers, adjuvants, stimuli-responsive constructs) can be bounded by toxicological priors and exposure scenarios, reducing late-stage attrition. For environmental engineers, we connect catalyst and sorbent design to remediation performance with tech-to-impact traceability via technology readiness levels (TRLs) and LCAs (EU Publications Office, 2017). Finally, for data/AI researchers, we outline benchmark desiderata—cross-domain datasets, active-learning loops, uncertainty calibration, and interpretability—that matter for safe deployment in wet labs and pilot plants (Tom *et al.*, 2024).

Concretely, we argue that converging AI + quantum + automation enables closed-loop, sustainability-aware discovery with measurable gains. AI—including foundation models and agentic planners—generates diverse, constrained candidates and proposes informative experiments; automation executes compact measurement campaigns and returns real-time signals to update the models; quantum and high-fidelity physics provide corrections where classical surrogates struggle, such as correlated surfaces, spin-dependent steps, and non-adiabatic processes (Paudel *et al.*, 2022; Weidman *et al.*, 2024). Sustainability metrics (e.g., LCA midpoints and USEtox-based toxicity factors) then become first-class objectives rather than downstream audits, reframing discovery from “can we make it and will it work?” to “should we make it, and how do we make it safest and cleanest?”—a necessary shift for catalysis (high selectivity with low embodied energy), therapeutics (efficacy with minimized off-target toxicity), and remediation (capture or destruction with low secondary burden) (Nizam *et al.*, 2021).

Evidence that this convergence is real is accumulating at the AI↔automation interface. Large-scale model-driven candidate generation—followed by robotic validation—has shown that thousands to millions of plausible materials can be prioritized and that a meaningful fraction can be synthesized within days to weeks, demonstrating a scalable feedback loop of

algorithmic proposals → robotic synthesis/characterization → model updating (Tom *et al.*, 2024). Such self-driving laboratory architectures bring two additional advantages: safety (by isolating operators from hazardous steps) and provenance (by enforcing structured data capture), both critical when moving toward regulated domains like therapeutics and environmental remediation (Tom *et al.*, 2024).

What is new in this review relative to prior work is cross-domain integration plus a focus on deployment-oriented metrics that enable apples-to-apples comparisons across discovery pipelines. We emphasize three classes of “hard” metrics. First, energy and time: report sample throughput (e.g., new materials per day), the number of optimization steps saved by active learning, and wall-time or kWh per closed-loop cycle. Second, eco-toxicity and safety: incorporate USEtox-compatible toxicity characterization factors, solvent/precursor hazard scores, and uncertainty ranges alongside performance metrics. Third, deployment readiness: map progress using TRLs from concept (TRL 1–2) to lab validation (TRL 3–4), prototype (TRL 5–6), and pilot/field contexts (TRL 7–8), culminating in proven systems (TRL 9) (EU Publications Office, 2017). Systematically reporting these metrics clarifies where quantum corrections add value, when foundation models eliminate unnecessary experiments, and how automation improves safety and reproducibility, while revealing bottlenecks that matter for translation beyond the lab (Paudel *et al.*, 2022; Weidman *et al.*, 2024; Pyzer-Knapp *et al.*, 2025).

This integration also reframes discovery as a multi-objective control problem under real-world constraints. From the AI perspective, active-learning policies must balance exploitation of promising regions with exploration under uncertainty while respecting sustainability penalties. From the automation perspective, experimental design must prioritize robust, information-dense measurements that are compatible with hazard-reduced chemistries and energy-lean processing. From the quantum perspective, hybrid quantum-classical stacks should be targeted to the highest-leverage physics gaps, supplying corrections only where they materially change down-stream decisions (Paudel *et al.*, 2022; Weidman *et al.*, 2024). The shared language across these perspectives is decision-theoretic: expected improvement tempered by environmental cost and operational risk.

Finally, we set expectations for the remainder of the article. We will detail agentic AI and foundation-model tooling for hypothesis generation and experiment planning; identify where quantum algorithms most usefully augment classical and ML models; describe a “minimum viable loop” for automation (design → synthesize → characterize → update) with uncertainty-aware decision rules; and operationalize sustainability with LCA/USEtox-compatible utilities and TRL-based

reporting. The aim is not to celebrate isolated advances but to specify practices and benchmarks that move candidates from computational screening to self-driving laboratories and into pilot-scale reactors, delivery systems, and remediation units—measured by accuracy and yield, yes, but equally by energy, speed, and minimized environmental burden (Tom *et al.*, 2024; Pyzer-Knapp *et al.*, 2025; EU Publications Office, 2017; Nizam *et al.*, 2021).

2. Foundations: Data, Representations, and Benchmarks

Modern discovery stacks sit on data. What makes today different is not just volume but heterogeneity: molecules and reactions from synthesis logs; crystals, surfaces, and defects from electronic-structure workflows; and biological and environmental measurements that connect a material to efficacy, safety, and fate. A useful way to organize this landscape is by modality. Molecular sources include linear notations such as SMILES and SELFIES (robust to invalid strings), graph formalisms that treat atoms as nodes and bonds as edges, fragment vocabularies for generative models, and reaction corpora with atom mapping, yields, and conditions (Krenn *et al.*, 2020; Coley *et al.*, 2019; Lowe, 2017). Materials sources capture periodicity and locality: crystal graphs with lattice/periodic images, slab models for surfaces and adsorption, and explicit defect/supercell enumerations; these are now standard in resources such as Matbench, JARVIS-DFT, OQMD, and the Open Catalyst Project where adsorption structures and relaxation trajectories define learning targets (Dunn *et al.*, 2020; Choudhary & Tavazza, 2020; Saal *et al.*, 2013; Chanussot *et al.*, 2021). Bio/environmental sources add the application layer: protein targets and binding pockets (e.g., PDBbind families), omics-derived features for mechanism-aware models, pollutant classes (PFAS, pesticides, pharmaceuticals), and kinetic/partitioning parameters relevant to fate and transport (Liu *et al.*, 2017; Huang *et al.*, 2021; NORMAN Network, 2021). The unifying theme is that discovery increasingly requires joined-up datasets: a catalyst is not just a bulk crystal; a therapeutic is not just a SMILES string; and a remediation agent is not just an adsorption energy—each must connect upstream structure to downstream performance and risk.

A FAIR (Findable, Accessible, Interoperable, Reusable) pipeline is essential to make these modalities usable across labs and over time. In practice, FAIR begins with curation (removing duplicates, harmonizing units and conditions), standardization (canonicalization of structures, charge states, isotopes; reaction templates; adsorption-site labeling), and licensing/provenance so downstream users know what they can share and reproduce (Wilkinson *et al.*, 2016; Boeckhout *et al.*, 2018). In chemistry and reactions, community efforts such as the Open Reaction Database (ORD) define schemas for reactants, reagents, solvents, catalysts, and outcomes, plus instrument metadata; analogous moves in

materials include Matbench task cards and JARVIS task definitions that specify input assumptions and calculation protocols (Coley *et al.*, 2021; Dunn *et al.*, 2020; Choudhary & Tavazza, 2020). For bio/env data, Therapeutics Data Commons (TDC) and EPA/ECOTOX-style repositories emphasize dataset cards, intended-use statements, and split recommendations to reduce leakage and enable apples-to-apples comparisons (Huang *et al.*, 2021). Provenance is not a formality: recording software versions, pseudopotentials/force fields, and experimental IDs is crucial for traceability and for de-duplicating near-identical entries that would otherwise inflate performance.

Representations translate raw data into machine-usable form. Graphs dominate for molecules and crystals: message-passing neural networks (MPNNs), CGCNN, SchNet, DimeNet(+), and equivariant models encode local chemical environments and, for materials, periodic boundary conditions (Gilmer *et al.*, 2017; Xie & Grossman, 2018; Schütt *et al.*, 2018; Klicpera *et al.*, 2020; Batzner *et al.*, 2022). 3D fields (voxel or continuous) capture electron density, electrostatic potential, or pocket geometry for docking and binding-affinity tasks, with SE(3)-equivariant networks bridging graphs and fields. Descriptors such as SOAP and MBTR remain powerful baselines when data are scarce, enabling kernel and linear models with strong inductive bias (Bartók *et al.*, 2013; Huo & Rupp, 2017). Finally, learned embeddings—from language-like tokenizers (SMILES/SELFIES) to contrastive or masked-prediction pretraining on crystals and surfaces—provide transferable features that can be fine-tuned for property, synthesis, or control tasks (Krenn *et al.*, 2020; Park *et al.*, 2023). Across choices, two principles help: (i) align representation with task physics (e.g., include periodic images for adsorption; encode chirality and 3D geometry for docking), and (ii) prefer equivariance when target properties transform predictably under rotations/translations.

Benchmarks are the community's contract: what do we claim to measure? For catalysis, open benchmarks focus on adsorption energies, surface relaxation, and reaction barriers as surrogates for turnover (Chanussot *et al.*, 2021). OC20/OC22 provide tens of millions of DFT single-point and relaxation labels across adsorbates and surfaces, with tasks ranging from initial-to-relaxed energy prediction (IS2RE) to force inference; gaps remain in explicitly measuring TOF/TON, stability under cycling, sintering/poisoning resistance, and support effects, which are critical for deployment but scarce in standardized, ML-ready form (Tran *et al.*, 2023). For drug discovery, public suites cover docking, binding affinity (ΔG via FEP/TI as higher-fidelity targets), and ADMET endpoints; common resources include PDBbind/CASF for structure-based tasks and TDC/MoleculeNet for ligand-based ADMET and safety (Liu *et al.*, 2017; Su *et al.*,

2020; Huang *et al.*, 2021; Wu *et al.*, 2018). Persistent gaps include off-target risk quantification at scale and clinical translatability proxies; scaffold splits help but do not fully address temporal drift and chemical novelty. For environmental technologies, datasets emphasize degradation kinetics, selectivity against co-contaminants, reusability under cycling, and leaching/secondary pollution; however, labels are heterogeneous (conditions, matrices, detection limits), and cross-study harmonization is a bottleneck (NORMAN Network, 2021; Nizam *et al.*, 2021). Across domains, deployment metrics—energy per synthesis step, yield per hazard score, TRL progression—are underrepresented but essential.

Reproducibility and uncertainty are the load-bearing beams of credible benchmarking. Uncertainty has two main flavors: aleatoric (data noise/irreducible) and epistemic (model uncertainty due to limited data). Practical toolkits include MC dropout, deep ensembles, and evidential regression for continuous properties; calibration metrics such as expected calibration error (ECE) and conformal prediction to produce valid prediction sets at a chosen error rate (Lakshminarayanan *et al.*, 2017; Guo *et al.*, 2017; Angelopoulos & Bates, 2023). In discovery loops, uncertainty must drive decisions: active learning should sample where epistemic uncertainty is high and penalize candidates with high eco-toxicity or safety risk. Reproducibility also hinges on splits. Time- or scaffold-based splits better reflect prospective performance than random splits in molecular tasks; in materials, composition- or structure- holdouts better emulate discovering new chemistries or prototypes than i.i.d. splits (Yang *et al.*, 2019; Dunn *et al.*, 2020). Leakage pitfalls include near-duplicates (e.g., salts/tautomers counted twice), train-test overlap via pretraining, and shared synthetic routes or DFT parameters sneaking across splits. Strong baselines must therefore publish deduplication rules, split hashes, and data cards describing what is—and is not—being measured.

A recurring practical challenge is license and use-rights. Public-private boundaries matter: industrial reaction notebooks, HTS screens, and pilot-plant logs often outclass public sets in scale and realism but come with restrictive licenses. Where possible, hybrid strategies—federated learning, secure enclaves, and synthetic data generated under privacy constraints—help bridge the gap without leaking proprietary content (Vepakomma *et al.*, 2018). Even in fully public settings, explicit SPDX-style license tags and provenance chains (what changed, when, by whom, with which tool) are necessary to make models reusable beyond their original authors.

Putting these pieces together, a discovery-ready dataset typically requires: (1) canonical structures (e.g., inchikeys for molecules; standardized CIFs and slab builders for surfaces), (2) task-ready targets (energies,

barriers, ΔG , ADMET, kinetic constants) with units and conditions, (3) negative and “boring” examples (failed syntheses, inactive compounds, poisoned catalysts) to prevent success bias, (4) uncertainty estimates (replicates

or model-derived) and recommended splits, and (5) a data card stating scope, intended use, caveats, and ethical/environmental constraints.

Table 1: Dataset and benchmark landscape

Domain	Representative datasets/benchmarks	Size (approx.)	License/Access	Primary tasks	Known caveats
Catalysis (surfaces)	Open Catalyst OC20/OC22	10^7 – 10^8 labels (DFT single-points/relaxations)	Open	Adsorption/relaxation energies; forces; initial→relaxed prediction	TOF/TON not explicit; limited sintering/poisoning labels
Materials (bulk)	Matbench; JARVIS-DFT; OQMD	10^5 – 10^6 entries	Open	Property prediction (bandgap, formation energy, elasticity)	Varying DFT settings; composition/structure leakage risks
Drug (structure-based)	PDBbind; CASF	10^3 – 10^4 complexes	Mixed academic	Docking/ranking; binding ΔG	Crystal packing bias; limited kinetics
Drug (ligand/ADMET)	TDC; MoleculeNet	10^4 – 10^6 molecules	Open	ADMET classification/regression	Scaffold/time splits essential; assay heterogeneity
Environmental	ECOTOX-like; NORMAN SusDat	10^3 – 10^5 chemicals/records	Open	Toxicity, degradation, partitioning	Condition heterogeneity; matrix effects; sparse negatives
Reactions	Open Reaction Database (ORD)	10^5 – 10^6 reactions (growing)	Open	Yield prediction; retrosynthesis; condition optimization	Incomplete atom mapping; yield/reporting bias

This table maps the core dataset/benchmark landscape across catalysis, materials, drug discovery, environmental science, and reactions, summarizing typical sizes, licenses, and primary tasks. Use it to pick fit-for-purpose data (e.g., adsorption vs. ΔG vs. ADMET) and to anticipate evaluation style (scaffold/time splits; composition/structure holdouts).

3. AI Methods for Inverse Design

Inverse design frames discovery as “specify the properties, then search the space of structures and processes that realize them.” Practically, that means three moving parts: (i) predictive models that map structure/process $\rightarrow$ properties with calibrated uncertainty; (ii) generative models that propose valid, synthesizable candidates while respecting safety and cost; and (iii) optimization loops that decide what to try next under multiple objectives and constraints. Around this engine sit physics-based priors, interpretability tools, and synthesis planners that turn virtual designs into routes in the real world.

Predictive models. Graph neural networks (GNNs) and message-passing networks remain the workhorses for molecules and materials because they encode local chemical environments and (for crystals) periodicity (Gilmer *et al.*, 2017; Xie & Grossman, 2018). Equivariant architectures further respect the symmetries of 3D space, improving data efficiency and force/energy consistency for atoms-in-materials and adsorbates-on-surfaces (Klicpera *et al.*, 2020; Batzner *et al.*, 2022). Transformer variants extend beyond sequences to graphs

and 3D point clouds, offering global receptive fields, strong transfer, and multi-task head sharing (Ying *et al.*, 2021; Park *et al.*, 2023). In low-data regimes common to catalysis or niche ADMET tasks, multi-task learning, transfer pretraining, and few-shot/adaptation strategies (e.g., metric-based or gradient-based meta-learning) can stabilize training and recover performance (Caruana, 1997; Altae-Tran *et al.*, 2017; Dunn *et al.*, 2020). Across these models, uncertainty estimation (ensembles, evidential heads) and calibration (ECE) are not luxuries—they are control knobs for safe decision-making in closed loops (Lakshminarayanan *et al.*, 2017; Guo *et al.*, 2017).

Generative design. Variational autoencoders learn smooth latent spaces that support gradient navigation toward property-optimized structures; junction-tree and fragment-aware VAEs help enforce chemical validity and scaffold realism (Gómez-Bombarelli *et al.*, 2018; Jin *et al.*, 2018). Normalizing flows provide exact likelihoods and invertible maps, making them attractive for conditioning on desired properties and for density-based active learning (Papamakarios *et al.*, 2017). Diffusion models, now state-of-the-art in molecular and 3D generative tasks, excel at capturing multi-modal structure distributions (Ho *et al.*, 2020; Hoogeboom *et al.*, 2022). For materials, diffusion/flow models over graphs and fractional coordinates increasingly handle periodicity, defects, and adsorption geometries. Evolutionary strategies and reinforcement learning (RL) remain competitive where objectives are discontinuous or the action space includes

edits to routes/process conditions rather than only structures (Brown *et al.*, 2019; Zhou *et al.*, 2019). Crucially, constraints must ride inside the generator: synthesizability proxies (e.g., SA-score, SCScore), cost/availability of precursors, and toxicity/hazard screens prevent “pretty but impossible” designs from saturating the loop (Ertl & Schuffenhauer, 2009; Coley *et al.*, 2018; Huang *et al.*, 2021).

Optimization loops. Inverse design becomes powerful when prediction and generation are coupled by Bayesian optimization (BO) and active learning (AL). BO balances exploitation and exploration using acquisition functions (e.g., expected improvement) that can be generalized to vector-valued, noisy objectives (Snoek *et al.*, 2012; Frazier, 2018). AL schedules the next experiment or computation to maximally reduce epistemic uncertainty, often with batch selection and diversity penalties to avoid redundancy (Settles, 2009). Because real discovery is never single-objective, multi-objective optimization tracks Pareto fronts over activity/efficacy, stability, cost, and toxicity (Deb *et al.*, 2002). A practical tactic is to scalarize with time-varying weights early (fast screening), then switch to Pareto-efficient selection once the knees of the curve emerge; another is to include risk terms (e.g., penalties for hazardous reagents or high energy/CO₂ per cycle) so the loop remains sustainability-aware.

Physics-informed AI. Data alone rarely constrain the search, so physics acts as a scaffold. “Hard” constraints (stoichiometry, charge/spin, symmetry, boundary conditions) can be baked into architectures—e.g., SE(3)-equivariance, periodic padding, or conservation layers—while “soft” constraints enter loss functions as regularizers (Schütt *et al.*, 2018; Klicpera *et al.*, 2020). Differentiable physics and operator-learning surrogates let models backpropagate through PDE solvers or tight-binding/DFT-like approximations, providing gradients that reflect real invariants and reducing spurious optima (Raissi *et al.*, 2019; Pfaff *et al.*, 2021). In catalysis, surrogate models pre-screen adsorption and barrier energies before expensive relaxations, and in therapeutics, differentiable docking or learned scoring functions provide physics-aware signals that stabilize generative training (Chanussot *et al.*, 2021; Su *et al.*, 2020). The art is choosing fidelity wisely: hybrids that call high-fidelity physics only where it will likely change rank order tend to dominate end-to-end throughput.

Interpretability. Inverse design must explain what it is doing, especially when choices have safety implications. Attention maps in graph transformers, substructure saliency for message-passing, and post-hoc attributors such as SHAP values can highlight which atoms, fragments, sites, or process features drive predictions (Ying *et al.*, 2019; McCloskey *et al.*, 2019; Lundberg & Lee, 2017). Yet interpretability is not causality. To avoid “Clever Hans” shortcuts (e.g.,

spurious correlations from assay conditions), teams increasingly combine attribution with counterfactuals (minimal edits that flip predictions) and with causal-inference ideas such as invariance testing across environments or interventions (Schölkopf *et al.*, 2021). For regulators and scale-up partners, interpretable decision records—what was proposed, why it was chosen, and what evidence supported it—are as important as raw scores.

Synthesis-aware design. A candidate is only as good as the route that makes it. In small-molecule drug design, retrosynthesis planners (template-based, template-free, and mixed) produce tree- or graph-structured routes subject to constraints on reagents, number of steps, yield priors, and cost (Segler *et al.*, 2018; Coley *et al.*, 2019; Schwaller *et al.*, 2020). For materials and nanomaterials, route planning means selecting precursors, solvents, temperatures/pressures, and time/atmosphere windows consistent with phase diagrams, safety, and scale (Dunn *et al.*, 2020). Embedding route constraints inside the design loop helps avoid dead-ends and leverages procurement reality (availability, pricing, hazard classes). A practical heuristic is two-stage generation: first, generate route-feasible candidates (filters: SA/SCScore, precursor lists, solvent classes), then fine-tune structure and process jointly using BO/AL while tracking embodied energy and hazard metrics. For remediation media and catalysts, include aging, sintering/poisoning resistance, and regeneration steps as part of the objective so the loop “sees” lifecycle costs, not only fresh performance.

Putting it all together (a typical closed loop). Start from a seed pool (legacy compounds, known materials, scaffold libraries, or enumerated adsorption geometries). Train a calibrated predictor with multi-task heads (property + uncertainty). Use a generator (diffusion/flow/VAE/RL) conditioned on target vectors and route constraints to propose a diverse batch. Run multi-objective BO to pick experiments that maximize acquisition while spreading along the Pareto front, with explicit diversity and safety terms. Execute the batch—via computation (DFT/FEP/TI) or SDLs—log metadata and failures, and update models (including uncertainty). Repeat, occasionally inserting high-fidelity physics where the model is uncertain but decisions are sensitive. Terminate when Pareto improvements saturate or when a candidate clears predefined gates (e.g., TRL-aligned criteria).

Common failure modes and mitigations. (i) Mode collapse in generators → enforce diversity with determinantal point processes, nucleus sampling, or diversity-aware acquisitions. (ii) Data leakage → use scaffold/time/composition-aware splits; hash and publish splits; quarantine pretraining overlaps. (iii) Reward hacking in RL/generative settings → add realism constraints (SA/SCScore, route cost), human-in-the-loop vetoes, and physics-based validators. (iv) Overconfident

predictors → ensembles + temperature scaling + conformal prediction to produce valid coverage. (v) Non-stationary objectives (e.g., updated toxicity assays) →

adopt time-aware evaluation and reweighting; re-validate surrogates when protocols change.

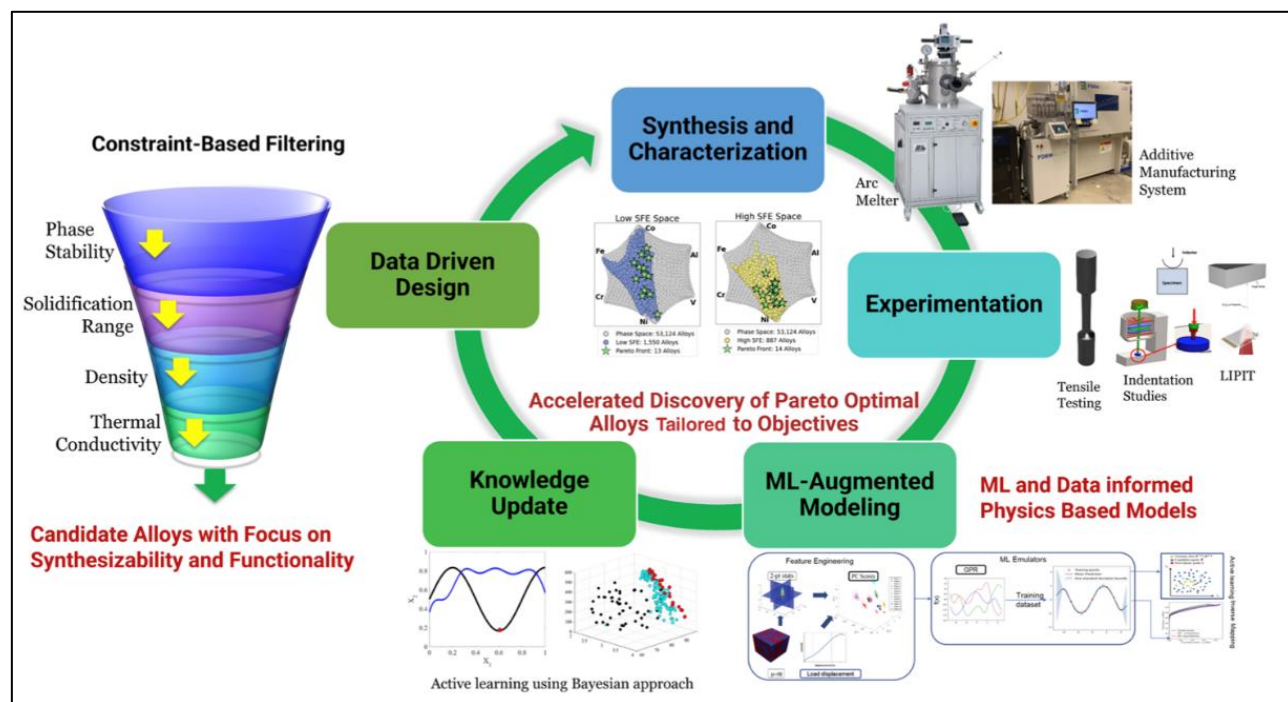


Figure 1. Closed-loop AL/BO inverse-design workflow

A calibrated predictor and a conditional generator propose route-feasible candidates under synthesizability and toxicity constraints. A multi-objective Bayesian optimizer selects batches that balance efficacy/activity, stability, cost, and hazard, routing high-uncertainty items to higher-fidelity physics or experiments. Results feedback to update both predictor and generator, advancing the Pareto front each cycle.

4. Quantum & Atomistic Engines

Design loops rise or fall with the physics underneath them. The practical rule is simple: use the cheapest model that is trustworthy for the decision at hand, and escalate fidelity only where it could flip the rank order among candidates. In this section, we assemble a working stack—from Kohn–Sham density functional theory (DFT) and its many-body corrections, to phonons and reaction barriers, solvation and interfacial realism, free-energy methods, machine-learning interatomic potentials (MLIPs), near-term quantum computing, and finally multi-fidelity workflows that couple these tools to maximize throughput without losing accuracy (Mardirossian & Head-Gordon, 2017).

Electronic-structure stack. Ground-state structures and energetics are typically obtained with Kohn–Sham DFT along Perdew’s “Jacob’s ladder”: LDA → GGA (e.g., PBE) → meta-GGA (e.g., SCAN) → hybrids (e.g., HSE, PBE0) → double hybrids, with DFT+U for localized d/f shells and dispersion

corrections as needed. Climbing rungs trades cost for accuracy; method choice should be justified per target (bonding type, correlation, charge transfer). For excited states, GW corrects quasiparticle levels and Bethe–Salpeter (BSE) captures excitons—essential for photocatalysis and optoelectronic screening. Reaction pathways rely on the nudged elastic band (NEB) family—especially the climbing-image variant—to locate minimum-energy paths and saddle points (Henkelman *et al.*, 2000). Lattice dynamics via phonons (harmonic or anharmonic) probe dynamical stability, finite-temperature free energies, and thermal transport (Togo, 2023). A robust recipe for catalysis or solid-state screening is: (i) relax structures with a meta-GGA or screened hybrid on a subset, (ii) map key barriers with NEB, (iii) compute phonons on shortlisted candidates, and (iv) apply GW/BSE only when spectra or level alignment could change decisions (Mardirossian & Head-Gordon, 2017; Henkelman *et al.*, 2000; Togo, 2023).

Beyond-DFT and ML interatomic potentials. Where DFT is marginal—strong correlation, dispersion-dominated binding, multi-reference pockets—one can stitch in CCSD(T) or other high-level references on fragments to Δ -correct DFT energies. For long times and large systems, MLIPs deliver near-DFT forces at orders-of-magnitude lower cost. Equivariant models such as NequIP encode rotational and permutational symmetries, improving data efficiency and stability for nanosecond-scale molecular dynamics (MD), rare-event sampling,

and thermal transport. In practice: curate a diverse active-learning set (including strained/defected/TS-like geometries), train an equivariant MLIP with uncertainty monitoring, validate on barriers/elastic constants/phonons, then run long MD or kinetic Monte Carlo to connect mechanisms to rates (Batzner *et al.*, 2022).

Solvation and interfaces. Real chemistry happens in solvents and at interfaces. Three modeling tiers are common. (1) Implicit solvation adds a continuum dielectric—fast and often sufficient for trends. (2) Explicit solvent (shells or slabs) captures hydrogen bonding, structuring, and entropic effects; combine with umbrella sampling when barriers matter. (3) Electrochemistry demands constant-potential (grand-canonical) DFT, where the electron chemical potential is controlled and the double layer is represented; constant-charge results can mislead when compared to experiments at fixed potential. Protein–ligand or protein–surface problems require protonation/ionic-strength realism and careful ensemble choice (Melander *et al.*, 2024).

Free-energy methods. Decisions hinge on free energies (ΔG), not just electronic energies: solvation, binding, selectivity, phase stability, and reaction equilibria. In molecular settings, free-energy perturbation (FEP) and thermodynamic integration (TI) yield absolute or relative ΔG ; umbrella sampling + WHAM reconstructs potentials of mean force along collective variables. In condensed phases and at surfaces, constrained MD (blue-moon ensembles) and anharmonic corrections are used; MLIPs make the needed sampling affordable. Modern variance-reduction schemes—e.g., mapped reference potentials—accelerate convergence by shaping sampling toward high-variance regions. Budget these expensive calculations where rank order is tight (e.g., $\Delta\Delta G$ selectivity within tens of meV or ~ 1 kcal·mol⁻¹) and rely on calibrated surrogates elsewhere (Rizzi, Rehbein, Zeller, & Hummer, 2021).

Quantum computing (forward look). Near-term devices remain noisy and small, but variational quantum eigensolvers (VQE) and related hybrid methods have matured on strongly correlated fragments and model Hamiltonians. The credible near-term role in design loops is specialist oracle: invoke a mitigated VQE calculation only where classical surrogates disagree and decisions are sensitive (e.g., spin-crossover centers, multi-reference adsorbates). Progress in error mitigation (symmetry checks, zero-noise extrapolation, learned noise models) defines feasibility windows for such calls (Jiang, Sun, Shaydulin, Lubasch, & Liu, 2024).

Multi-fidelity workflows. The throughput multiplier is layering: couple fast but approximate models (descriptors, MLIPs, GGA-DFT) with high-accuracy corrections (hybrids, GW/BSE, CCSD(T),

explicit free energies) only where they are likely to change rank order. Two patterns dominate. Δ -learning uses sparse paired labels to learn the difference between cheap and expensive levels and applies it broadly. Adaptive routing with multi-fidelity Bayesian optimization/active learning sends a candidate to a higher rung when its value of information is high (i.e., uncertainty is large and the decision is sensitive); otherwise it stays on the cheap track. This approach yields higher Pareto throughput (activity/efficacy, stability, cost, toxicity) at fixed budget, and it integrates naturally with laboratory automation and safety filters. The Jacob's-ladder intuition helps communicate the idea: ascend only when necessary (Mardirossian & Head-Gordon, 2017).

Heterogeneous catalysis. Screen adsorption on key facets with GGA/meta-GGA; fit an equivariant MLIP for long-time coverage and site-disorder effects; escalate a handful of candidates to NEB for rate-limiting steps and to hybrids where charge transfer is delicate; if photophysics matters, compute GW/BSE on finalists; close with microkinetics to estimate TOF under operating conditions (Henkelman *et al.*, 2000; Togo, 2023).

Drug discovery. Use docking and learned scorers to prune; reserve FEP/TI for near-ties in $\Delta\Delta G$; include explicit solvent/ions for charged series and water networks; push flexible systems with MLIP-accelerated sampling; gate candidates by ADMET surrogates before costly physics (Rizzi *et al.*, 2021).

Environmental remediation. Prioritize binding/selectivity versus co-contaminants; simulate regeneration/aging (fouling, poisoning) and leaching in explicit solvent; verify framework stability by phonons and thermodynamics; route constant-potential electrochemical steps to grand-canonical DFT (Melander *et al.*, 2024).

Checks and common pitfalls. (i) Functional sensitivity: re-compute a subset (e.g., SCAN $\rightarrow$ HSE) to detect cancellations. (ii) Finite-size artifacts: ensure slab thickness, k-points, plane-wave cutoffs, and counterpoise checks are converged—and record them for provenance. (iii) Double counting in Δ -corrections: match geometries and dispersion schemes across levels. (iv) Electrochemical realism: avoid comparing constant-charge simulations to constant-potential experiments. (v) Free-energy convergence: monitor overlap and statistical inefficiency; use replica exchange or enhanced sampling if barriers are rough. (vi) MLIP brittleness: maintain an active-learning loop with disagreement triggers and validate on forces and curvatures (phonons), not only energies (Henkelman *et al.*, 2000; Togo, 2023; Melander *et al.*, 2024; Batzner *et al.*, 2022).

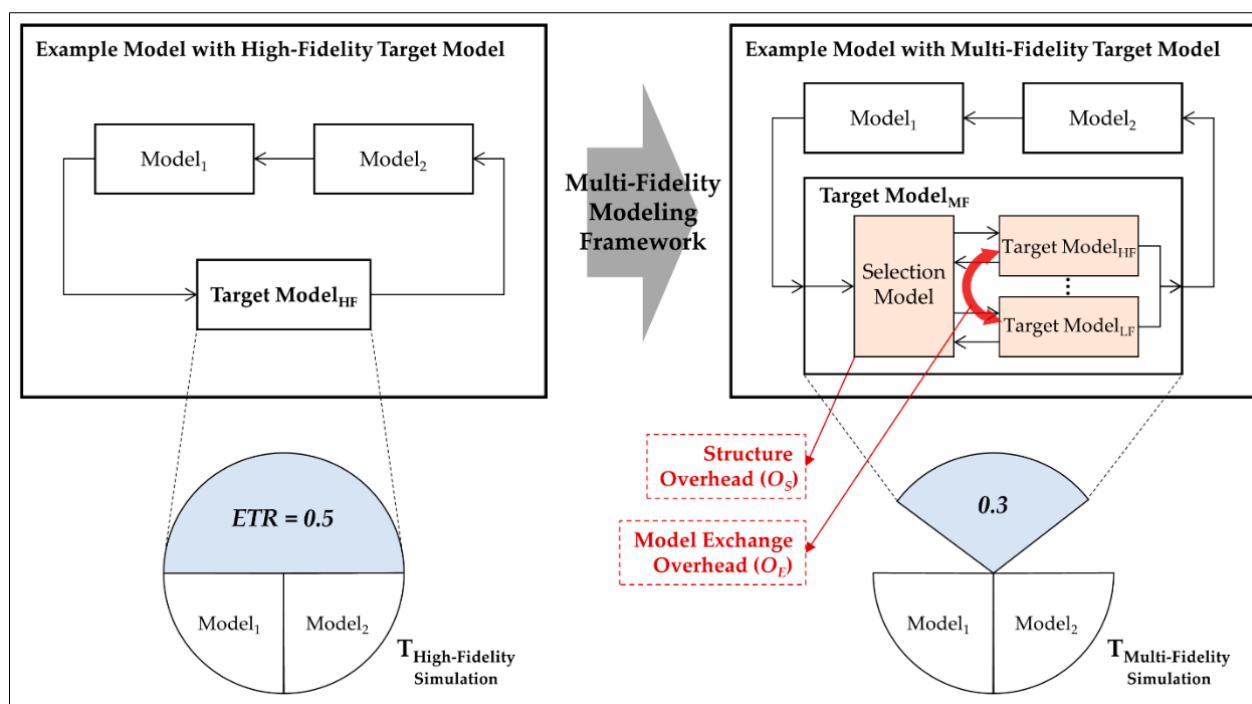


Figure 2. Multi-fidelity ladder for quantum & atomistic design loops

A low-fidelity branch (descriptors/MLIPs/GGA-DFT) screens quickly; a high-fidelity branch (hybrids, GW/BSE, CCSD(T), explicit free energy) validates only decision-sensitive points (Mardirossian & Head-Gordon, 2017). Information flows bidirectionally: sparse high-fidelity labels Δ -correct the cheap model; the cheap model supplies diverse proposals and uncertainty cues.

Adaptive routing via multi-fidelity Bayesian optimization/active learning escalates samples when value of information is high—balancing accuracy, cost, and throughput. Use this schematic to communicate fidelity-escalation policies and to justify compute allocation across rungs.

5. Materials Platforms & Design Spaces

Modern inverse design is only as powerful as the design space it explores. Here we organize practical platforms—nanostructures, frameworks and membranes, 2D materials and defect chemistries, and bio-hybrids—and outline knobs that move performance, the stability/scalability traps to watch, and how to couple these spaces to closed-loop optimization.

Discrete nanoparticles offer high surface-to-volume ratios and tunable coordination environments; quantum dots add size-quantized electronic states; core-shells and heterostructures decouple functions (e.g., light absorption vs. charge separation) while single-atom catalysts (SACs) maximize metal atom efficiency and uniformity of active sites (Wu & Yang, 2020; Chen *et al.*, 2021). Key knobs include composition (alloying, dopants), size/shape (facets, edges), support/ligand identity, and defect density. For photocatalysis or optical

delivery, quantum-confined dots (e.g., chalcogenides, perovskites) permit bandgap tuning across $\sim 1\text{--}3$ eV by radius control, while surface passivation governs non-radiative losses and colloidal stability (Li & Zeng, 2019). For thermal and electro-catalysis, sub-10 nm particles balance activity with sinter-resistance; SACs on nitrogen-doped carbons or oxide supports offer high turnover with reduced noble-metal loading, provided the anchoring coordination is robust under operating temperatures and redox swings (Chen *et al.*, 2021). In drug delivery, inorganic cores (Au, SiO₂, Fe₃O₄) can be shaped and coated for photothermal, imaging, or magnetic targeting functions, but surface corona formation and RES (reticuloendothelial system) uptake must be managed through ligand chemistry (Anselmo & Mitragotri, 2019).

Metal-organic frameworks (MOFs), covalent organic frameworks (COFs), and zeolites provide crystalline, modular scaffolds whose pore size, topology, and chemistry can be systematically varied (Diercks & Yaghi, 2017; Waller *et al.*, 2019). In catalysis and separations, hierarchical porosity (micro-meso) alleviates diffusion limits, while post-synthetic modification introduces catalytic motifs or hydrophobic/hydrophilic balance for complex feeds (Waller *et al.*, 2019). For mixed-matrix membranes (MMMs), dispersing MOF/COF/zeolite fillers in polymers can beat the permeability-selectivity trade-off by creating preferential pathways; success depends on interfacial compatibility and filler percolation (Pérez-Reyes *et al.*, 2021). Stability is platform-dependent: carboxylate-linked MOFs may hydrolyze; Zr-based nodes and imine-to- β -ketoenamine COFs improve water/thermal resistance; zeolites excel at hydrothermal

stability but are compositionally less flexible (Diercks & Yaghi, 2017; Waller *et al.*, 2019). For remediation, MOFs/COFs bearing chelating groups capture PFAS/heavy metals, but regeneration, fouling, and leaching define viability; embedding in robust MMMs mitigates particle loss and eases module integration (Pérez-Reyes *et al.*, 2021).

Transition-metal dichalcogenides (TMDs) such as MoS₂/WS₂, hexagonal BN, and doped graphene expose edge and basal-plane sites whose defect/strain chemistry can be engineered for catalysis, sensing, or separation (Voiry *et al.*, 2018; Zhao *et al.*, 2020). In hydrogen evolution, for instance, basal planes are inert while edge sulfur vacancies and 1T' phases activate sites; heteroatom doping (N, B, P) in graphene tunes adsorption energies and electron density, creating ORR/OER/CO₂RR-relevant ensembles (Voiry *et al.*, 2018; Zhao *et al.*, 2020). For membranes, stacked graphene oxide or MXene laminates produce angstrom-scale channels for ion sieving; swelling control and oxidation state stability are the limiting factors (Ding *et al.*, 2020). In optics, 2D excitons yield strong light–matter coupling but suffer photobleaching under high flux; encapsulation and defect passivation extend lifetimes (Zhao *et al.*, 2020). Defect engineering must balance activity with structural fragility: vacancy-rich lattices can reconstruct, and dopants may segregate under potential/temperature cycling (Voiry *et al.*, 2018).

Bio-hybrids. Enzyme–nanoparticle conjugates and nanozymes marry catalytic specificity with nanomaterial robustness. Enzymes immobilized on porous oxides, carbons, or MOFs gain thermal and solvent tolerance; mass transfer through mesopores and retention of active-site orientation are the design levers (Li *et al.*, 2018). Nanozymes—nanostructured oxides, metals, or carbon allotropes with enzyme-like kinetics—offer low-cost, scalable alternatives for peroxidase/oxidase-mimicking reactions in sensing and therapeutics, though substrate specificity and in vivo compatibility remain challenges (Wang *et al.*, 2020). For drug delivery, lipid nanoparticles (LNPs) and polymeric carriers (PLGA, PEGylated blocks) control biodistribution via size (~60–150 nm), surface charge, and ligand targeting; endosomal escape and payload stability are the gating mechanisms, and batch-to-batch reproducibility under GMP constraints is critical for translation (Kulkarni *et al.*, 2021; Anselmo & Mitragotri, 2019). Hybrid constructs—e.g., enzyme-loaded MOF shells or stimuli-responsive polymer–inorganic composites—enable cascade catalysis or on-demand release but add interfacial failure modes (Li *et al.*, 2018).

Across platforms, stability loss mechanisms usually dominate lifetime and cost. In nanoparticle catalysts, sintering and Ostwald ripening coarsen size distributions; supports that anchor single atoms/clusters

through strong metal–support interactions and defect-rich carbons delay coalescence (Chen *et al.*, 2021). In MOFs/COFs/MMMs, hydrolysis, linker oxidation, and polymer plasticization degrade performance; crosslinkers and robust nodes (e.g., Zr, Ti) raise tolerance (Waller *et al.*, 2019; Pérez-Reyes *et al.*, 2021). For 2D materials and quantum dots, photobleaching, photo-oxidation, and ligand desorption under illumination and heat degrade optoelectronic response; inorganic shells and short-chain, multidentate ligands help (Li & Zeng, 2019). In bio-hybrids, corona formation, enzyme denaturation, and carrier aggregation alter targeting and kinetics; protein-repellent coatings and mild immobilization chemistries mitigate these (Anselmo & Mitragotri, 2019; Li *et al.*, 2018).

Scale strategy strongly influences costs and footprints. Batch wet-chemistry remains dominant for colloids and COFs/MOFs but struggles with heat/mass transfer uniformity; continuous-flow microreactors and millifluidic systems improve control and reproducibility for nanoparticles and QDs, reduce solvent use, and integrate inline analytics for closed-loop control (Khan *et al.*, 2021). Mechanochemical synthesis (ball milling) avoids bulk solvents and can access otherwise difficult linkages in MOFs/COFs; solvent-lean microwave, photochemical, and supercritical CO₂ routes likewise support green chemistry goals when evaluated over full life-cycles (Friščić *et al.*, 2020). For membranes and MMM modules, roll-to-roll casting and phase-inversion lines dominate CAPEX decisions; filler alignment and dispersion uniformity are the hidden bottlenecks (Pérez-Reyes *et al.*, 2021). Across all spaces, integrating LCA metrics (energy per kg, solvent hazard scores, regeneration cycles) into optimization closes the loop between performance and sustainability.

First, define the minimum viable manifold of knobs that control the property of interest—e.g., particle size and ligand field for plasmonic heating; linker identity and node valency for MOF adsorption; defect type and areal density for TMD catalysis; or N/P ratio and lipid identity for LNP transfection. Second, choose representations that tie directly to those knobs (e.g., graph features for coordination, SOAP/MBTR for local environments, slit-pore descriptors for membranes, or learned embeddings for carrier composition), and bind them to constraints that reflect synthesis reality (temperature/solvent windows, precursor availability, GMP/biocompatibility). Third, encode failure modes as penalties or objectives—sintering propensity, hydrolytic stability, photobleaching rates, protein corona scores—so the optimizer “sees” lifetime and safety as early as activity. Finally, route stability testing into the loop: aging protocols, cycling under realistic feeds, and regeneration tests are as important as headline selectivity or turnover.

Table 2: Design-space catalog

Platform	Tunable knobs	Typical property ranges/examples	Synthesis/scale constraints
Nanoparticles / QDs / core–shells / SACs	Size (1–50 nm), shape/facet, alloying/dopants, ligand/support, single-atom coordination	Bandgap ~1–3 eV (QDs); TOF↑ with facet control; SAC M–N ₄ sites for ORR/CO ₂ RR	Sintering/Ostwald; ligand desorption; continuous-flow for narrow dispersity; support anchoring for SACs
MO / COFs / zeolites	Linker/node chemistry, topology, pore size/functionalization, hierarchical porosity	BET 500–6000 m ² g ^{−1} ; selective adsorption/separation; catalytic site installation	Hydrolysis/oxidation; particle shedding in MMMs; interfacial compatibility with polymers
2D (TMDs, h-BN, doped graphene)	Phase (2H↔1T'), vacancy density, heteroatom doping, strain, stacking	Edge/vacancy sites for HER/ORR; angstrom-channels for ion sieving	Swelling/oxidation; defect reconstruction; scalable exfoliation/CVD; laminate stability
Bio-hybrids (enzymes, nanozymes, carriers)	Enzyme loading/orientation, pore size, nanozyme composition, carrier size/charge/ligand	k _{cat} /K _M tuning; peroxidase-like activity; targeted delivery & controlled release	Corona formation; denaturation; GMP reproducibility; immunogenicity; solvent and pH windows

6. Application Domain I — Catalysis

Catalysis underpins low-carbon fuels, green chemicals, and polymer circularity. We focus on thermo-, electro-/photo-electro-, and photocatalysis across CO₂ reduction (CO₂RR), oxygen reduction/evolution (ORR/OER), nitrogen reduction (NRR), selective oxidations, and plastic upcycling—domains where AI-guided screening, quantum/atomistic validation, and microkinetics now run as one loop from idea to reactor (Motagamwala & Dumesic, 2020; Leonzio, 2024).

In heterogeneous catalysis, high-throughput DFT and learned surrogates map adsorption/barrier landscapes, while Brønsted–Evans–Polanyi (BEP) relations tie thermodynamics to kinetics. Microkinetic models then convert elementary energetics into rates and selectivities under realistic feeds, enabling volcano analyses and coverage effects. Practically: (i) pretrain surrogates on open and in-house slabs; (ii) run active learning/Bayesian optimization (AL/BO) to sample sites (facets, steps, defects, ensembles); (iii) escalate high-value points to higher fidelity (hybrids, explicit solvation, constant-potential DFT for electrocatalysis); and (iv) select batch experiments that maximize expected improvement in turnover frequency or selectivity subject to stability and cost (Chanussot *et al.*, 2021; Baz, Comas-Vives, & López, 2021; Göttl *et al.*, 2022).

- **Thermo-catalysis.** Rates reflect barriers governed by scaling and BEP; microkinetics plus microreactor data guide reactor choice (fixed/packed beds, fluidized or slurry systems). Dehydrogenation/oxidation selectivity often emerges from tuned bifunctional ensembles and acid–base balance. Plastic upcycling combines hydrogenolysis and selective C–C scission with coke management (Motagamwala & Dumesic, 2020).

- **Electrocatalysis.** Device-relevant metrics require gas-diffusion or MEA flow cells that manage mass transport, carbonate chemistry, and ohmic losses. Constant-

potential modeling clarifies Tafel slopes and rate orders; catalyst–ionomer–membrane architectures co-determine local fields and water management (Baz *et al.*, 2021; Lin *et al.*, 2022).

- **Photocatalysis/photo-electrochemistry.** Light absorption, charge separation, and interfacial transfer dominate; photo-corrosion and photobleaching drive lifetime. PV-coupled CO₂RR or redox-mediated architectures can decouple intermittency from reaction conditions while preserving selectivity (Motagamwala & Dumesic, 2020).

Report, at minimum: turnover frequency (TOF) and turnover number (TON); selectivity (molar/carbon basis); overpotential (η) for electro/photo steps; Faradaic efficiency (FE) and partial current density at specified cell voltages; stability with hour counts and drift criteria; and catalyst cost and earth-abundance proxies. For CO₂RR, add single-pass carbon efficiency, energy efficiency, and flow/MEA performance at ≥ 200 mA·cm^{−2} if claiming device readiness; recent reports of sustained C₂⁺ at ampere-scale current densities define today's benchmark bar (Leonzio, 2024; Lin *et al.*, 2022; Chen, Wang, Li, & Chen, 2024).

Dominant routes include coking, poisoning (S/Cl/P/halides), sintering/Ostwald ripening, dissolution/leaching (electrochemical), phase reconstruction, and support corrosion. Materials levers are strong metal–support interactions, defect-anchored single-atom sites, corrosion-resistant supports, and alloying to resist halides; process levers are periodic regeneration (oxidative burn-off/reduction), potential pulsing, electrolyte impurity control, and operation inside phase-stable windows. Durability claims should pair accelerated stress tests with long-hold runs (≥ 100 h) and post-mortems (TEM/XRD/XPS/ICP) that separate sintering from poisoning or leaching (Forzatti, 1999; Anekwe, Li, Salako, & Zhang, 2025; Pham *et al.*, 2025).

” BEP relations transform reaction energies into activation energies, enabling fast kinetic surrogates; generalized forms handle multi-step networks and site specificity. Microkinetics converts those into coverage-dependent rates/selectivities, revealing Sabatier volcano structure and identifying which adsorption-energy tweak would move a surface toward the peak without sacrificing stability (Götl *et al.*, 2022; Yang *et al.*, 2024; Motagamwala & Dumesic, 2020).

Case study flow (CO₂RR): AL → DFT → cell benchmarking.

- Seed & learn. Begin with Cu-rich, Ag-decorated, and defect-rich carbon motifs; train a surrogate on adsorption energies and key barriers from open OC datasets and in-house slabs. Use active learning to explore steps, grain boundaries, and heteroatom ensembles; treat the interface with constant-potential DFT where fields matter (Chanussot *et al.*, 2021; Cheng, Luo, & Cheng, 2022).
- Down-select via microkinetics. Convert energetics into predicted partial currents and selectivities across CO₂/H₂O/CO activities; penalize dissolution or reconstruction windows at target potentials/pH (Baz *et al.*, 2021).
- Validate physics. Escalate finalists to hybrid DFT for charge-transfer-sensitive steps; include explicit water/ionomer when trends hinge on local fields.
- Device-level test. Fabricate MEA/flow-cell electrodes; benchmark FE(C₂⁺), partial current

density at $\geq 200 \text{ mA} \cdot \text{cm}^{-2}$, energy efficiency, and drift over 24–100 h; compare to state-of-the-art cells reporting $\text{A} \cdot \text{cm}^{-2} \text{ C}_2^+$ production (Lin *et al.*, 2022; Chen *et al.*, 2024).

- Close the loop. Feed positives and negatives back to the surrogate; update BO to advance the Pareto front (selectivity–stability–cost). When ranks are ambiguous, request new high-fidelity labels (e.g., explicit-solvation free energies) (Götl *et al.*, 2022).

After clearing device bars, reactor models size mass/heat transfer and balance-of-plant. For thermocatalysis, packed-/fluidized-bed or slurry models constrain particle size, gradients, and pressure drop; for electrolyzers, gas-diffusion electrodes, CO₂ delivery, carbonate management, and membrane/ionomer durability dominate. Techno-economic analysis (TEA) ties productivity and lifetime to leveled cost, with sensitivity to electricity price, current density, FE, and stability. Report earth-abundance metrics to avoid breakthroughs that cannot scale (Leonzio, 2024; Lin *et al.*, 2022).

Encode durability proxies—sintering propensity, halide/sulfur binding, leaching windows—directly into the objective so AL/BO avoids fragile optima. Track regeneration efficacy and post-mortem signatures to drive causal fixes rather than parametric overfitting (Forzatti, 1999; Pham *et al.*, 2025).

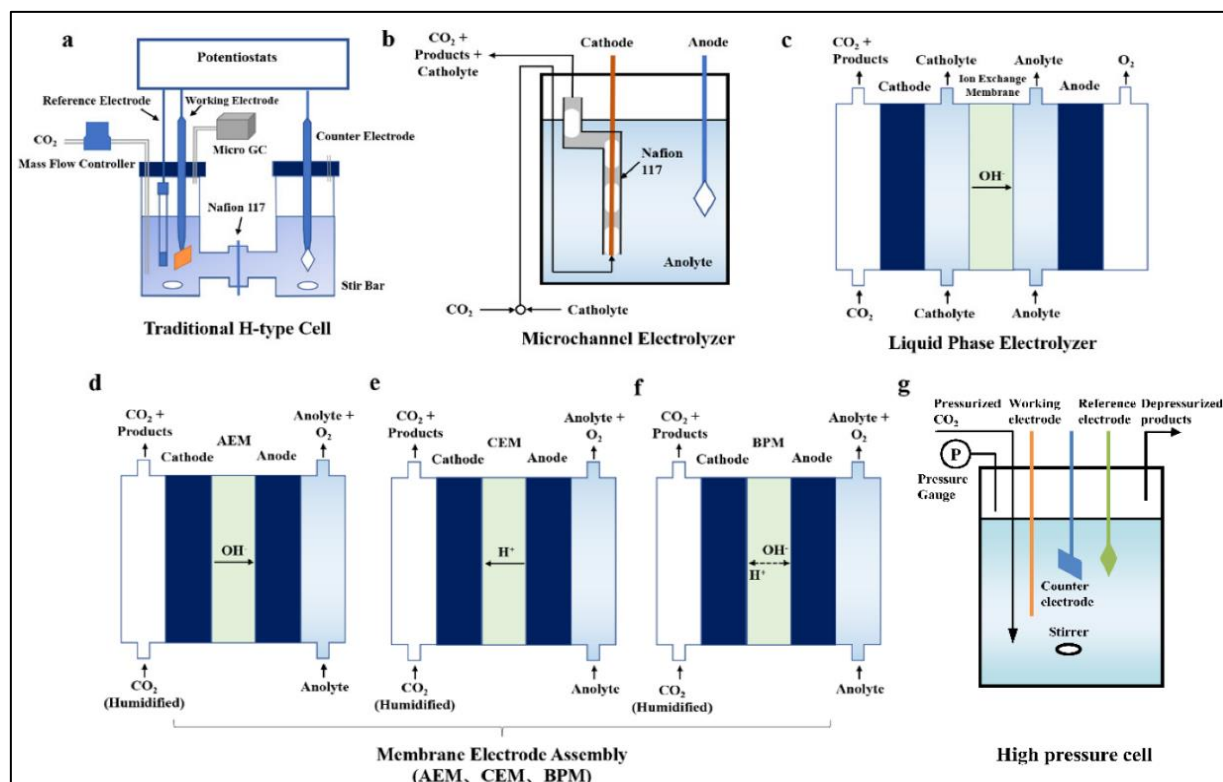


Figure 3: CO₂ electroreduction cells used for device-level benchmarking (real schematic).

Diagram comparing H-cell, microchannel/flow-cell, and MEA architectures commonly used to evaluate CO₂RR catalysts at increasing mass-transport and current-density demands. Highlights where device-relevant metrics (partial current density, FE, energy efficiency) are obtained and why flow/MEA cells are essential for $\geq 200 \text{ mA} \cdot \text{cm}^{-2}$ operation. Place this figure adjacent to your CO₂RR case study to anchor “AL → DFT → cell” transitions in real hardware

7. Application Domain II — Drug Discovery

Drug discovery now spans a wide target space—enzymes, GPCRs, ion channels, protein–protein interfaces (PPIs), and nucleic-acid targets—and multiple modalities, from small molecules and peptides/peptidomimetics to oligonucleotides and nano-enabled delivery systems. Since 2017, the biggest shift has been toward integrated funnels that couple structure/dynamics of targets with generative chemistry, calibrated scoring, physics-based free energies, and early ADMET triage, all guided by uncertainty and synthesis constraints (Ekins *et al.*, 2019; York *et al.*, 2023).

Target space & modality choice. Enzymes remain tractable because pocket geometry constrains design, but PPIs and RNA/DNA motifs are increasingly addressable via hot-spot mapping, fragment merging, macrocycles, and peptide/peptidomimetic scaffolds. Structure sources blend crystallography/cryogenic EM with AlphaFold-class models and molecular dynamics to reveal cryptic or induced-fit pockets that reshape SAR. Modality should be chosen for *biophysics + distribution*: small molecules for intracellular enzymology; macrocycles/peptides for shallow PPIs; siRNA/ASO payloads when gene-level control is needed; and nanocarriers when permeability, stability, or tissue access limit efficacy (Mehta *et al.*, 2023; Kim *et al.*, 2023).

Screening funnels. A robust end-to-end loop is staged and data-calibrated:

- Generative ideation. Constrained generation (scaffolds, pharmacophores, synthesizability) proposes diverse chemotypes; synthesis scores and building-block availability gate feasibility.
- Docking + rescoring. Docking is treated as a pose *proposer*; ML rescoring (graph/transformer models) provides calibrated ΔG estimates with uncertainties.
- Physics refinement. Alchemical FEP/TI or enhanced-sampling PMFs resolve near-ties ($\sim 1 \text{ kcal} \cdot \text{mol}^{-1}$) and water-network edge cases; these steps are reserved for shortlists to control cost.
- ADMET triage. Permeability/solubility, metabolic stability and CYP liabilities, hERG risk, and reactive substructure flags reduce late attrition.

- Make–test–analyze. Prospective assay data (biophysics, enzymology/cell, early PK) are fed back to re-train scorers and generators under uncertainty, closing the loop (York *et al.*, 2023; Ekins *et al.*, 2019).

Nanocarriers as part of the design problem. Lipid nanoparticles (LNPs), liposomes, polymeric nanoparticles/micelles, dendrimers, and inorganic cores (gold, silica, magnetic) extend the accessible pharmacology by controlling biodistribution, release, and stability. Design knobs include size ($\approx 60\text{--}150 \text{ nm}$ for systemic use), surface charge, PEG density, ligand display, core/shell chemistry, and endosomal-escape motifs. Failure modes—protein corona, aggregation, complement activation, and immunogenicity—must be screened early, and carrier choice must be co-optimized with payload potency and PK; a vehicle cannot rescue weak pharmacology (Alshawwa *et al.*, 2022; Mehta *et al.*, 2023; Kim *et al.*, 2023).

Safety & translation. Early off-target prediction reduces rework: hERG (K_v11.1) risk models combine ligand-based and structure-informed features; CYP inhibition/induction predictors anticipate DDIs; broader in-silico panels (kinases, GPCRs, ion channels) and transcriptomic/phenotypic profiling flag liabilities. Lessons since 2019: (i) hERG models improve with assay-consistent labels and protein context; (ii) CYP models benefit from multimodal inputs and domain adaptation to novel chemistry; and (iii) nanocarrier safety hinges on orthogonal tests (hemolysis, complement, macrophage uptake) and PBPK that accounts for corona-driven sequestration (Garrido *et al.*, 2020; Weiser *et al.*, 2023; Kim *et al.*, 2023).

CMC and scalability. Translation stalls when chemistry, manufacturing, and controls (CMC) lag behind discovery. For small molecules this means convergent routes, impurity fate maps, and solid-form control; for nanotherapeutics it requires scalable mixing (e.g., microfluidic LNP production), in-process analytics, lot-to-lot reproducibility, and GMP-ready CQAs (size/PDI/encapsulation, stability). Lock a DoE early to avoid “one-off” lab formulations that fail at scale (Liu *et al.*, 2021; Agha *et al.*, 2023).

Clinical & regulatory notes; digital twins (forward look). A credible preclinical package documents target engagement, exposure–response, GLP tox (including ion-channel safety), and a control strategy linking process to product quality (ALCOA+ data integrity). Digital twins/virtual patients are emerging to pressure-test dose, inclusion criteria, and responder enrichment before Phase trials—but demand external validation and bias checks (An, 2022; Alsalloum *et al.*, 2024).

Putting it together: End-to-end molecule + nanocarrier loop.

- Choose the modality per target biology and delivery constraints (enzyme vs PPI vs RNA; free-drug vs carrier).
- Constrained generation → docking/ML- ΔG → FEP/TI for shortlists to resolve rank ties and waters.
- ADMET triage (permeability, metabolic stability, hERG/CYP/off-targets, reactivity flags) with calibrated uncertainty.
- Carrier co-design only when PK/access demands it; screen corona/immune interactions and manufacturability.
- Prospective calibration via make-test-analyze cycles; lock scalable routes (for LNPs, microfluidic mixing and lyophilization).
- Preclinical to IND with GLP tox, stability-indicating methods, and a CMC control strategy linking CQAs to clinical material. Across all steps, treat uncertainty as a first-class signal to route expensive physics and experiments where they change decisions (York *et al.*, 2023; Mehta *et al.*, 2023; Liu *et al.*, 2021).

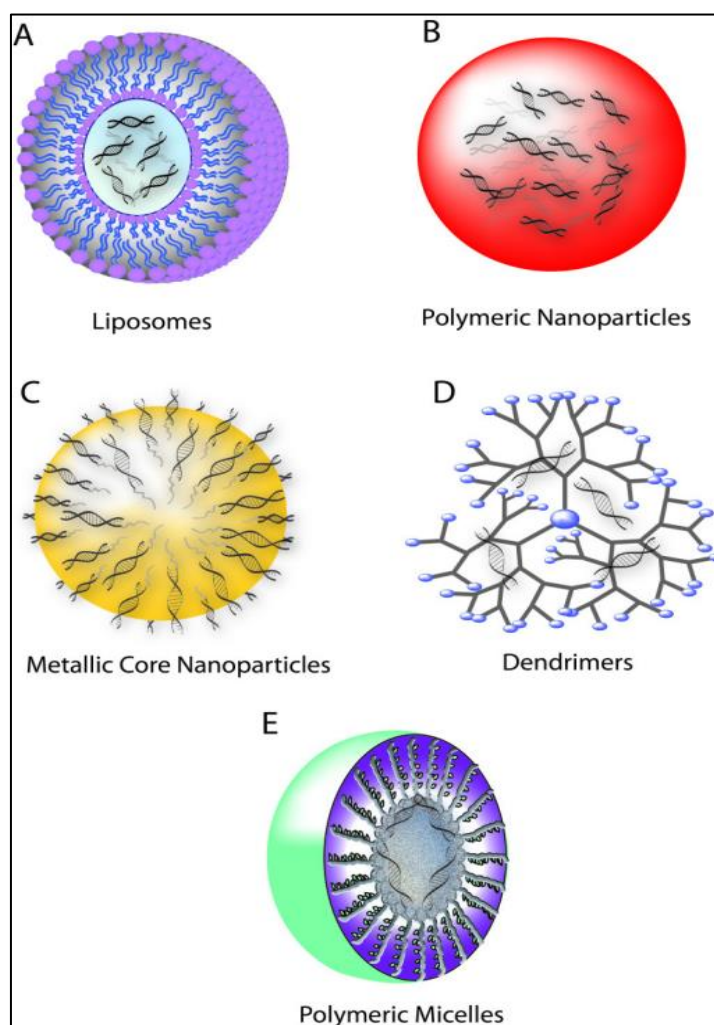


Figure 4: Representative nanocarrier classes for therapeutic delivery.

Panel illustrating liposomes, polymeric nanoparticles, metallic-core nanoparticles, and dendrimers—highlighting how carrier architecture affects loading and release. Use this to anchor the “molecule + nanocarrier” decision: when potency is adequate but PK, stability, or targeting limit efficacy, the design space expands from ligand chemistry to carrier CQAs. Place near the nanocarrier paragraph and CMC notes to signal the translation pathway from formulation choice to GMP-ready attributes.

8. Application Domain III — Environmental Remediation

Industrial effluents and diffuse pollution now present mixed contaminant streams—dyes with high chromophoric stability, heavy metals (As, Pb, Cd, Hg) lacking degradation pathways, PFAS with extreme fluorinated persistence, pharmaceuticals with bioactivity at $\text{ng} \cdot \text{L}^{-1}$, microplastics as colloidal carriers of sorbed toxics, VOCs impacting air and water, and NO_x/SO_x in flue gas–water couplings. Effective remediation therefore matches pollutant class → mechanism →

material under realistic matrices (ionic strength, NOM, co-ions) and tracks not only removal but byproducts, durability, and cost/energy (Crini & Lichtfouse, 2019; Wang, DeWitt, Higgins, & Cousins, 2017). Mechanisms and where they win.

- Adsorption/ion exchange excels for hydrophobic dyes, many APIs, and metals (outer-/inner-sphere complexation), offering modular columns and straightforward regeneration. Selectivity hinges on surface functionality (e.g., $-\text{COOH}/-\text{NH}_2$), microporosity, and competing anions; for arsenic, iron (oxy)hydroxide and Zr-based sorbents form inner-sphere complexes that resist desorption (Crini & Lichtfouse, 2019).
- Photocatalysis & advanced oxidation (AOPs) mineralize organics via $\bullet\text{OH}$ or $\text{SO}_4\bullet^-$, with TiO_2 and $\text{g-C}_3\text{N}_4$ as workhorses; heterojunctions and dopants shift band positions and suppress recombination, while persulfate activation broadens pH windows and matrix tolerance (Huang, He, & Zhong, 2019; Wang & Wang, 2020).
- Electrocatalytic degradation (EAOPs) leverages reactive oxygen on BDD or dimensionally stable anodes; paired with electrosorption or Fe/Cu mediators, EAOPs handle persistent pharmaceuticals and PFAS precursors, and integrate readily into modular reactors (Garcia-Segura, Ocon, & Chong, 2018; Qiao, Guo, & Sun, 2023).
- Membranes deliver phase barriers (MF/UF/NF/RO) and reactive separations (photocatalytic or adsorptive layers). They excel for microplastics and metals but face fouling, concentration polarization, and retentate handling—hence hybrid trains (adsorption $\rightarrow$ NF/RO $\rightarrow$ AOP) are common (Qiu, Zhang, & Zhao, 2019).
- Materials palette and tuning.
- MOFs/COFs/zeolites. MOFs/COFs offer designer pore chemistry and high surface areas; sulfonated/aminated nodes capture dyes and metals, while Zr-based nodes tolerate water/oxidants. Zeolites bring hydrothermal stability and ion exchange capacity for NH_4^+ /heavy metals (Dong, Tu, & Zheng, 2020).
- Carbons (biochar/AC/graphene derivatives). Surface oxygen/nitrogen groups and π - π domains bind dyes and APIs; doped carbons introduce Lewis basicity and redox mediation. Biochars deliver low-cost capacity with ash/alkali side effects; activation and acid washing tailor selectivity (Tran, Ok, & Sik, 2017).
- Semiconductors. TiO_2 (UV) and $\text{g-C}_3\text{N}_4$ (visible) form Z-schemes or S-schemes with oxides/sulfides to widen spectral response and suppress recombination; perovskites and

bimetallics (e.g., Ag–Cu, Fe–Ni) tune adsorption and radical generation at interfaces (Wang & Wang, 2020; Huang *et al.*, 2019).

- Bio-sorbents and hybrids. Chitosan, alginate, and protein-functional materials chelate metals and bind dyes; immobilizing enzymes in MOFs or on oxides adds biocatalytic steps. Stability and microbial growth control are the design constraints (Crini & Lichtfouse, 2019).
- Durability, leaching, and secondary pollution. Capacity alone is not sufficiency. Assess: (i) leaching of metals/ligands from MOFs or nanoparticle supports; (ii) byproduct toxicity—AOP intermediates, short-chain PFAS from precursor oxidation; (iii) photobleaching of sensitizers and corrosion under EAOPs; and (iv) regeneration routes (thermal/solvent/pH swing, electro-regeneration) with minimal capacity fade. Report cycle life (≥ 5 –10 cycles), mass balance (parent + intermediates + mineralization), and post-mortem analytics (XPS/ICP/TEM) to deconvolute mechanisms (Dong *et al.*, 2020; Garcia-Segura *et al.*, 2018; Wang & Wang, 2020).
- PFAS, microplastics, and metals: three stress tests.
- PFAS. C–F persistence defeats conventional AOPs; the field favors separate-and-destroy: high-affinity capture (ion-exchange resins, fluorophilic sorbents) followed by destructive steps (electrochemical oxidation/reduction, UV-sulfite, plasma). Avoid partial defluorination that yields equally persistent/toxic fragments; fluoride balance and targeted HRMS for precursors are required (Wang *et al.*, 2017; Qiao *et al.*, 2023).
- Microplastics. Size-fractionated removal by coagulation–flocculation, DAF, and membranes is effective; aging increases polarity and adsorption of co-pollutants, calling for upstream carbon or hybrid AOPs to avoid contaminant-rich concentrates (Qiu *et al.*, 2019).
- Metals (As, Pb, Cd, Hg). Favor inner-sphere complexation and redox stabilization (e.g., As(V) precipitation, Hg(II) complexation) on iron/zinc oxides, thiolated carbons, or MOF linkers; design around competing anions (PO_4^{3-} , SO_4^{2-}) and pH windows (Crini & Lichtfouse, 2019; Dong *et al.*, 2020).

From bench to field. Laboratory wins must survive matrix effects (hardness, NOM, alkalinity) and fouling (biofouling, scaling). Fieldable trains often adopt adsorption $\rightarrow$ membrane $\rightarrow$ AOP sequences to separate, concentrate, and finally destroy, or EAOP polishing post-RO to treat retentate. Pilot $\rightarrow$ municipal scaling requires skid-mounted modules, backwash/clean-in-place SOPs, and energy/cost accounting ($\text{kWh}\cdot\text{m}^{-3}$; $\text{\$}\cdot\text{m}^{-3}$) including regenerant disposal. TEA/LCA should

report media $\text{\$}\cdot\text{kg}^{-1}$, bed volumes to breakthrough, specific energy per log removal, and sludge/retentate handling to avoid burden shifting (Garcia-Segura *et al.*, 2018; Qiu *et al.*, 2019).

Putting the decision tree to work. Start with *contaminant identity + matrix*: (1) Hydrophobic dye in high-NOM water? $\rightarrow$ hydrophobic carbon or COF + low-dose UV/PS AOP; KPI = capacity ($\text{mg}\cdot\text{g}^{-1}$), pseudo-second-order rate, color removal, energy per m^3 . (2) As(V)/Pb(II) in groundwater? $\rightarrow$ Fe/Zr oxide or thiolated

carbon with pH-swing regeneration; KPI = breakthrough bed volumes at regulatory limit, selectivity over $\text{HCO}_3^-/\text{SO}_4^{2-}$. (3) PFAS (C4–C8) in municipal influent? $\rightarrow$ ion-exchange or fluorophilic adsorbent + electrochemical destruction of regenerant; KPI = Σ PFAS and fluoride balance, energy per mmol F^- released. (4) Mixed pharmaceuticals ($\mu\text{g}\cdot\text{L}^{-1}$) $\rightarrow$ NF/RO + UV/ H_2O_2 or EAOP polishing; KPI = parent/intermediate toxicity and specific energy. (Crini & Lichtfouse, 2019; Wang *et al.*, 2017; Garcia-Segura *et al.*, 2018; Qiu *et al.*, 2019).

Table 3: Remediation KPIs with standardized units

KPI	Definition	Typical unit	Notes for reporting
Capacity	Sorbate uptake at equilibrium (per mass of sorbent)	$\text{mg}\cdot\text{g}^{-1}$	Report isotherm (Langmuir/Freundlich), temperature, pH, ionic strength
Rate constant	Apparent kinetic constant (pseudo-first/second order)	min^{-1} or $\text{g}\cdot\text{mg}^{-1}\cdot\text{min}^{-1}$	Provide reactor type (batch/column), film/mass-transfer limits
Selectivity	Preference vs. competing species	dimensionless (ratio) or %	Specify competitors and concentrations (e.g., SO_4^{2-} , Cl^- , NOM)
Removal	Fraction of parent removed	% or $\log_{10}$ reduction	Pair with mineralization (% TOC/fluoride balance) to avoid byproduct masking
Energy	Specific electrical/UV energy	$\text{kWh}\cdot\text{m}^{-3}$ or $\text{kJ}\cdot\text{g}^{-1}$ pollutant	Include duty cycle, electrode/UV efficiency, and matrix absorbance
Stability	Performance over reuse	cycles to 20% capacity loss	Include regeneration protocol, leachate analysis (ICP, LC-MS)
Pressure drop/Flux	Hydraulic performance (membranes/columns)	kPa ; $\text{L}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$	Report fouling control (CIP/backwash), temperature, crossflow
Cost	Media/reactor cost normalized	$\text{\$}\cdot\text{m}^{-3}$ treated	Include media lifetime, regenerant/disposal costs (avoids burden shifting)

This table standardizes what to report so results are comparable across materials and pilots. Pair removal with mineralization/toxicity to prevent green-washing via persistent intermediates. Express energy and **cost** on a per-volume basis, with media lifetime and waste handling. Always include matrix descriptors (NOM, hardness, pH, co-ions) and uncertainty to support scale-up decisions.

9. Closed-Loop Experimentation & Automation

Modern discovery programs increasingly hinge on closed-loop experimentation—hardware and software that plan, execute, analyze, and then decide the next experiments with minimal human intervention. The core ingredients are (i) high-throughput experimentation (HTE) and robotics to generate dense, high-quality data; (ii) active-learning planners that balance exploration and exploitation; and (iii) standards that let instruments and informatics talk reliably, with rich metadata and robust error handling. When these pieces click, self-driving workflows compress months of manual iteration into days while maintaining traceability and reproducibility (Tom *et al.*, 2024; Christensen *et al.*, 2021).

HTE & robotics. Microfluidic and mesofluidic platforms now serve as agile “workhorses” for synthesis, formulations, and materials processing. Arrays of microliter-scale reactors, controlled by syringe/pressure manifolds and LED modules, can sweep temperature,

residence time, stoichiometry, light intensity, and reagent identity with exquisite repeatability. Inline/online analytics—UV-Vis/IR probes, MS/ESI-MS, HPLC/UHPLC with autosamplers, Raman, DLS, even compact XRD—collapse feedback time from hours to minutes, so the planner learns on the fly rather than after a batch campaign. For solid-handling or membrane/coating work, collaborative robots and liquid handlers prepare libraries, while integrated balances, viscometers, tensiometers, and contact-angle tools measure key physical attributes. Critically, the physical layout should be “planner-aware”: parallel reactors feeding a shared analytics queue, with barcoded consumables and fail-safe waste routing to prevent cross-contamination (Shields *et al.*, 2021; Guo, Ranković, & Schwaller, 2023).

Active learning with sample-budget limits. Because experiments are costly, the planner must choose wisely. Bayesian optimization (BO) is the default engine for continuous conditions, using a surrogate (often a Gaussian process) with an acquisition function (e.g., expected improvement, upper confidence bound) that trades off exploitation (trying the current best neighborhood) and exploration (reducing uncertainty in poorly known regions). For reaction or formulation spaces that mix continuous and categorical choices (ligand, solvent, membrane polymer), specialized methods incorporate descriptors for the categorical

options to guide search in a “soft” space (Häse *et al.*, 2020). In multi-response settings—yield *and* selectivity, or flux *and* rejection—multi-objective BO pushes a Pareto front rather than a single optimum. Two pragmatic constraints dominate real labs: batching (proposing *N* experiments in parallel for a plate/run) and scheduling (coordinating reactors and analytics). Planners therefore co-optimize the *which* (next conditions) and the *when/where* (assignment to modules), so idle time is minimized and analytical queues do not become bottlenecks (Ruan *et al.*, 2022; Shields *et al.*, 2021).

Scheduling and orchestration. Self-driving campaigns are rarely single-instrument affairs. A robust loop models the lab as a set of resources (reactors, shakers, ovens, chromatographs) with capacities and service times. The planner emits action graphs (“dose A, heat 60 °C for 8 min, irradiate 455 nm, quench, inject HPLC”), while a scheduler enforces resource constraints, retries transient failures, and logs state transitions. Practical heuristics—like “hedged batches” that include a few exploitation points, a few uncertainty-reduction points, and a couple of controls—keep learning on track when measurements arrive asynchronously. For heterogeneous tasks (e.g., membrane casting + permeability testing), the loop may switch objectives between phases: first maximize film integrity/defect-free casting, then optimize permeability–selectivity under fixed casting conditions.

Interfacing instruments (APIs, schemas, and provenance). To be reliable and shareable, an autonomous loop needs standardized control and data. On the control side, SiLA 2 defines device services and command/response structures so pumps, valves, photoreactors, and balances expose consistent APIs—vital when swapping vendors or scaling up. On the data side, experiment schemas keep context intact: for small-molecule reactions, the Open Reaction Database (ORD) schema captures reagents, operations, conditions, and outcomes in a machine-readable way; for analytical results, AnIML-style containers record spectra/chromatograms with method metadata and calibration. Minimal yet FAIR (Findable, Accessible, Interoperable, Reusable) pipelines start with strict deduplication and standardization, unit harmonization (SI), and provenance (who/what/when/how), then append data-quality flags (outlier detection, sensor health) and licensing so data can be reused downstream (Kearnes *et al.*, 2021; SiLA Consortium, 2019).

Error handling and robustness. Automation fails; resilient loops plan for it. Distinguish hard faults (device offline, comms failure) from soft faults (pressure spike, detector saturation, out-of-spec peak). Hard faults trigger re-routing or safe shutdown; soft faults yield censored data—which the model should ingest with uncertainty rather than discarding. Calibration drift is mitigated by interleaving standards and references;

concept drift (chemistry changes as the loop moves) is handled with time-aware kernels or periodic re-initialization around newly discovered regions. Every loop should maintain audit trails (actions, firmware versions, calibrations, chem inventory lots) to make results reproducible.

Mini-vignette 1 — Autonomous photoredox discovery. A flow photoreactor with LED modules (blue/green) and HPLC-MS inline analytics explores catalyst, base, solvent, light intensity, residence time, and stoichiometry. The planner begins with a diverse seeding design, trains a GP on conversion/selectivity, and proposes batches via expected improvement under a sample-budget of 12–24 experiments per hour. When the GP’s predictive variance spikes—e.g., a new photocatalyst family—it triggers exploration; near promising basins, it switches to local exploitation and, for close contenders, escalates to physics-based checks (e.g., transient spectroscopy to confirm productive excited states). Over a day, the loop finds higher selectivity at lower light power by shifting to a different photocatalyst/solvent pair and slightly longer residence time—an outcome a purely manual DoE would likely miss (Shields *et al.*, 2021).

Mini-vignette 2 — Autonomous membrane formulation. A casting robot prepares polymer–additive–solvent blends on glass, controls drying temperature/humidity, then runs inline thickness and defect imaging; cured films are mounted in a mini-permeation skid to measure permeability and selectivity for water/organic or gas pairs. The planner first optimizes film integrity (defect rate < 1%) under manufacturing constraints (max solids %, drying ramp), then switches objective to the permeability–selectivity Pareto front, proposing formulations that respect viscosity windows (pumpability) and cost caps. A multi-reactor/one-analyzer scheduler parcels batches to keep the permeation skid saturated; the best formulations are re-cast at larger area to verify scale effects (Christensen *et al.*, 2021).

People and process. Autonomy amplifies expert time rather than replacing it. Chemists and engineers still define objective functions (what does “good” mean?), constraints (synthetic plausibility, safety, waste), stop rules (sufficient confidence or resource cap), and validation plans (orthogonal assays, out-of-distribution tests). Teams should adopt model cards and data sheets for experiments, documenting assumptions, training data, and known failure modes; this avoids over-claiming and eases tech transfer. Finally, governance matters: version-controlled recipes, CI tests for instrument drivers, and simulation sandboxes for planner updates prevent “bricking the lab” during software changes (Tom *et al.*, 2024).

10. Safety, Sustainability, and Life-Cycle Thinking

Designing catalysts, therapeutics, and remediation materials to be safe and sustainable by design means elevating environmental and human-health considerations to the same level as performance and cost. In practice, that requires three tight couplings: (i) green chemistry metrics that steer route and process choices while the chemistry is still malleable; (ii) life-cycle assessment (LCA) and circularity thinking that extend the boundary from flask to factory to fate; and (iii) nanosafety principles and regulatory alignment so scale-up does not create new risks. When these are embedded in closed-loop discovery, the optimizer no longer seeks “the best material,” but rather the best feasible material under sustainability and safety constraints (Sheldon, 2017, 2018; Jiménez-González *et al.*, 2011).

Green chemistry metrics that drive decisions. Three yardsticks consistently influence outcomes early enough to matter. Atom economy rewards transformations that place most of the reactant atoms into product, but it is blind to workup and solvent. The E-factor (kg waste per kg product) broadens the lens to every stream exiting the process; high E-factors typically correlate with poor economics and environmental load. Process Mass Intensity (PMI), now common in pharma and fine chemicals, goes further by counting all inputs—solvents, reagents, auxiliaries—per kilogram of product, making it sensitive to recovery and recycle strategies. Together, E-factor and PMI let teams trade yield against separations burden and solvent volumes; both can be forecast from route sketches and refined as unit operations are locked (Sheldon, 2017, 2018; Jiménez-González *et al.*, 2011; Benison *et al.*, 2022; Kekessie *et al.*, 2024). Because solvents often dominate PMI, solvent selection guides have become the largest early lever: moving from chlorinated or high-toxicity ethers toward alcohols, esters, water, or carbonates can lower PMI and hazard without sacrificing process windows, provided drying and separation energies are accounted for (Benison *et al.*, 2022; Kekessie *et al.*, 2024).

From gate-level metrics to cradle-level footprints. Route- and step-level indicators are necessary but not sufficient. LCA extends analysis to the full life cycle—cradle-to-gate (raw materials to factory gate), cradle-to-grave (through use and end-of-life), or cradle-to-cradle (with recovery loops). Following ISO 14040/14044, credible LCAs specify goal/scope, system boundaries, functional unit, inventory sources, impact methods (e.g., ReCiPe), and uncertainty/sensitivity analyses. For materials achieving the same function, it is routine to observe burden shifting: a composition with lower climate impact may worsen freshwater ecotoxicity or mineral depletion. Communicating these trade-offs with normalized, multi-category radar (spider) charts helps decision-makers select options consistent with project priorities and policy constraints (Hollberg & Ruth, 2021). In closed loops, fast LCA surrogates—fed by bills of materials, unit operations, and assumed energy

mixes—can provide impact vectors that sit alongside performance metrics so multi-objective optimization is truly performance $\times$ footprint, not performance first, footprint later (Hollberg & Ruth, 2021).

Circularity and life-cycle sustainability assessment. Circular design targets recyclability, recoverability, and benign end-of-life, not just lower footprints at manufacture. Life-cycle sustainability assessment (LCSA) integrates environmental LCA with life-cycle costing and social metrics (worker safety, supply risk, community impacts), recognizing that a low-carbon material relying on scarce or conflict-exposed inputs is not legitimately “sustainable.” Adding critical-minerals flags and recovery yields as constraints in the design loop helps avoid stranded breakthroughs that cannot be responsibly scaled (Finkbeiner *et al.*, 2020).

Nano-(eco)tox: design rules for particles and interfaces. At the nanoscale, hazard and exposure depend on size, shape/aspect ratio, dissolution kinetics, and surface chemistry—all of which evolve in real media as “coronas” of proteins or natural organic matter adsorb. This bio/eco-corona rewires particle identity, tuning aggregation, transport, uptake, and immune recognition. Robust evidence now shows trophic transfer in aquatic food webs and context-dependent toxicity that shifts with ionic strength and organic matter. Reproducible assessment therefore requires standardized reporting of hydrodynamic size distributions, number concentrations, zeta potential, dissolution, and corona composition in relevant media, alongside realistic exposure metrics (Drasler *et al.*, 2017; Abdolapur Monikh *et al.*, 2020; Liu *et al.*, 2023; Zhang *et al.*, 2024). For discovery, these factors can be encoded as penalties or constraints: avoiding high-aspect-ratio shapes in environmental applications unless fate data are strong; preferring coatings that remain stable across expected pH/ionic-strength windows; and prioritizing compositions with low bioaccumulation potential.

Standardization is catching up—use it. The OECD Working Party on Manufactured Nanomaterials has issued and updated test guidelines relevant to particles, including guidance for size distribution measurement, dispersion stability, dissolution/leaching, and adaptations of bioaccumulation and toxicokinetics studies to nanoforms. These harmonized protocols, along with sector guidance from EFSA for agri-food and environmental contexts, enable comparability and regulatory acceptance across labs. Treat them as design constraints, not just compliance hurdles: plan characterization and hazard screens that map one-to-one to these methods so datasets are reusable and auditable (Drasler *et al.*, 2017; Abdolapur Monikh *et al.*, 2020).

Regulatory landscape to design for—not around. In the EU, REACH now requires explicit nanoform identification and characterization in registrations, with data bridging allowed only when

scientifically justified. In the United States, FDA guidance clarifies how nanomaterials in drug products trigger additional CMC and risk-benefit considerations, and EPA uses TSCA mechanisms (including significant new use rules) to manage new nanoscale substances. Cross-cutting actions on persistent classes—e.g., expanding PFAS reporting and phase-out schedules—illustrate how surfactants, processing aids, or polymer additives can suddenly fall under stricter scrutiny, reshaping material choices and manufacturing routes (ECHA, 2020; FDA, 2022; EPA, 2024). Building regulatory readiness into optimization—by tagging candidates that would trigger special reporting or nanoform dossiers—prevents late-stage redesigns.

Operationalizing “safe & sustainable by design” in a loop. Practically, make sustainability a first-class objective. Train multi-task predictors to output performance (activity, selectivity, potency), processability (solubility windows, synthesis steps), and proxy hazard (e.g., aquatic toxicity tiers; hERG/CYP for drugs) with calibrated uncertainty; route low-confidence regions to higher-fidelity tests. Attach a solvent plan and PMI/E-factor snapshot to every candidate proposal so Bayesian optimization trades performance against mass intensity at acquisition time. Bundle a fast LCA vector with each recipe using scenario-appropriate energy and

transport inventories to expose climate, toxicity, and resource-use trade-offs during selection. For nano-enabled options, require OECD-compatible characterization and leaching/dispersion screens before escalating exposure scenarios, and treat regeneration and end-of-life as optimization targets (cycles to 20% capacity loss; recovery yields). Finally, document assumptions and data provenance model cards for predictors, data sheets for experiments, and change logs for PMI/LCA snapshots so sustainability claims are transparent and reproducible (Sheldon, 2018; Kearns *et al.*, 2021; Hollberg & Ruth, 2021).

Putting the spider chart to work. When three materials achieve the same functional unit (e.g., equal conversion or dose efficacy), compare normalized LCA categories on a spider chart beside KPI tables. A candidate that minimizes climate and fossil resource impacts might score worse on ecotoxicity due to metal leaching or solvent choices; another might excel on toxicity but depend on a scarce element. Selecting the Pareto-efficient option then becomes a policy-aware choice, not a single-metric race. Include uncertainty bands and sensitivity to hotspots (e.g., solvent recovery rate, electricity mix) so readers can see where additional data would change the decision (Hollberg & Ruth, 2021).

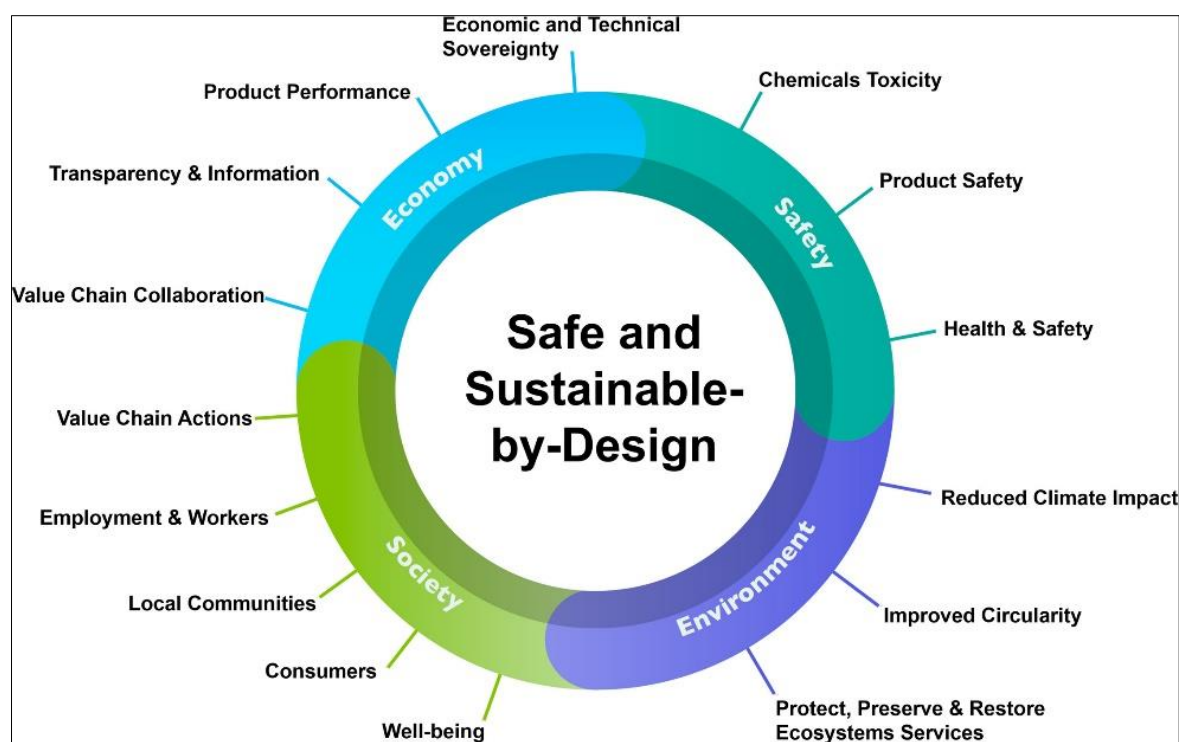


Figure 5: Comparative Life-Cycle Impact Radar for Function-Equivalent Materials

This section makes sustainability a first-class objective alongside performance, using atom economy, E-factor, and PMI to steer routes and solvent choices early. We extend the boundary with life-cycle assessment and circularity, comparing candidates on climate, toxicity, and resource use for the same function.

Nano-(eco)tox rules—size, surface, dissolution, and corona—are encoded as constraints, with standardized assays for reproducibility. Finally, we align designs with evolving regulations (REACH, FDA, EPA) so winners are deployable, not just publishable.

11. Cross-Cutting Physics & Theory

Across catalysis, therapeutics, and remediation, the same physical primitives govern behavior: how structure fixes electronic and vibrational states; how interfaces shape charge and mass exchange; how kinetics couples to transport under operando conditions; and how we pass information across time and length scales. Making these connections explicit lets AI and automation optimize what is physically achievable rather than what is merely convenient to compute (Motagamwala & Dumesic, 2020).

Structure $\rightarrow$ property. Point and extended defects, elastic strain, and quantum confinement remap bands, phonons, and dielectric response, thereby tuning reactivity, transport, and stability. Vacancies and antisites introduce mid-gap states and trap carriers, while grain boundaries host strained, chemically distinct ensembles that often become the true active sites in heterogeneous catalysis. In soft, polar semiconductors—most famously halide perovskites—strong electron–phonon coupling stabilizes large polarons that screen charged defects and renormalize effective masses; transport, recombination, and hot-carrier cooling therefore deviate from simple band pictures (Frost, 2017; Meggiolaro & De Angelis, 2020). Quantum confinement raises exciton binding energies in 2D materials and quantum dots, and modest biaxial strain can shift adsorption energies and band edges enough to move a catalyst along a Sabatier volcano or a bioelectronic sensor across a detection threshold (Ghosh, Zhou, & Wong, 2020). These links set the hard bounds for inverse design: microstructure and fields can flip trends for the same chemistry.

Interfacial phenomena. Real devices are typically interface-limited. At electrochemical solid–liquid boundaries, the electric double layer (EDL) sets the local potential drop, screening length, and solvent structure; specific adsorption, finite-size sterics, and the quantum capacitance of the electrode all modulate kinetics beyond classical Gouy–Chapman–Stern models (Baz, Comas-Vives, & López, 2021). Band alignment at aqueous interfaces depends not only on surface termination but also on hydration, surface dipoles, and pH-controlled protonation; explicit-solvent *ab initio* molecular dynamics combined with dielectric-continuum embedding has clarified how these factors pin absolute energy levels relevant to photocatalysis and sensing (Hörmann, Ambrosio, & Pasquarello, 2019). In nano-bio contexts, rapidly evolving solvation shells and protein or humic coronas rewrite particle identity within milliseconds, changing adhesion, charge-transfer pathways, and cellular uptake—one reason simulations and experiments must represent the intended medium (Liu, Zhang, & Lowry, 2023).

Kinetics and transport under operando. Rates arise from elementary steps embedded in fields and flows. Microkinetic models convert adsorption energies

and transition states into coverages, rates, and selectivities while honoring site balances and lateral interactions; coupling these to mass, heat, and charge transport explains why a surface that looks selective in UHV becomes diffusion-limited or potential-limited in a flow cell, membrane, or tissue (Motagamwala & Dumesic, 2020). In electrocatalysis, constant-potential kinetics should be solved self-consistently with the EDL to capture Tafel slopes, reaction orders, and buffering effects (Baz, Comas-Vives, & López, 2021). In separations and remediation, reaction–diffusion equations with Donnan partitioning describe depletion layers and selectivity in charged pores; in drug delivery, permeability and binding kinetics embed within PBPK-style tissue transport, turning *in vitro* potency into exposure–response (An, 2022).

Bridging time and length scales. No single method spans femtoseconds to hours or ångströms to millimeters, so we assemble a ladder. At the electronic/atomistic level, DFT (with +U or hybrids where localization matters) and GW/BSE for excitations resolve local chemistry and excitons; phonons and non-adiabatic couplings capture vibronic effects essential for polarons and hot carriers (Frost, 2017; Meggiolaro & De Angelis, 2020). Event-level kinetic Monte Carlo (kMC) then projects rare barrier-crossing events—surface reconstruction, sintering, defect migration—onto laboratory timescales using rate catalogs derived from atomistics, increasingly with time-dependent fields and open-system boundary conditions (Yang *et al.*, 2024). For soft matter and nano-bio interfaces, coarse-grained (CG) models and machine-learned CG force fields compress many-body atomistics into tractable beads while preserving thermodynamics and key kinetics, enabling studies of permeation, aggregation, and corona evolution at orders-of-magnitude lower cost (Durumeric, Vani, & Onuchic, 2023; Ye & Zhang, 2021). Finally, continuum models—reaction–diffusion–migration PDEs, phase-field for morphology, Nernst–Planck–Poisson or Darcy–Brinkman in porous media—deliver device-level fluxes, selectivity, and stability; parameters “flow up” from atomistic/kMC/CG, while boundary conditions “flow down” from reactors or physiology.

Information contracts and uncertainty. Multiscale workflows succeed when each level passes upward compact, meaningful summaries—barrier distributions, adsorption isotherms, diffusivities, partition coefficients—and requests from above only those boundary conditions it truly needs—fields, loadings, and chemical potentials. Two patterns are especially effective. First, train surrogates with guardrails: fast predictors (equivariant GNNs for energies, graph transformers for ΔG , neural CG forces) that respect symmetries, conservation laws, and correct far-field limits, and that emit calibrated uncertainty so active learning can request the most valuable new labels (Shields *et al.*, 2021). Second, use embedded coupling: microkinetics inside continuum solvers for reactors or

tissues; EDL/solvation models inside electrochemical kinetics; or kMC event rates coupled to diffusion or elasticity so local depletion and stress feed back on event frequencies (Yang *et al.*, 2024). Treating uncertainty as a first-class quantity turns these stacks into decision engines rather than just simulators.

Implications for inverse design. Structure–property maps, interfacial band diagrams, and coupled kinetics–transport define feasible regions before optimization begins. Polaronic screening can render certain dopants benign—or catastrophic—depending on dielectric response (Frost, 2017). EDL structure and specific adsorption can move a volcano summit across electrolytes, so a “best” electrocatalyst at one pH fails at another (Baz, Comas-Vives, & López, 2021). Diffusion–reaction coupling can turn a top-ranked surface into a mass-transfer-limited one in flow cells or porous scaffolds; in membranes, phase-field evolution and kMC growth rules predict pore connectivity and fouling propensity; in drug delivery, CG predictions of mucus penetration or corona-mediated opsonization propagate into PBPK exposure and therapeutic index (An, 2022; Durumeric, Vani, & Onuchic, 2023). Embedding these mechanisms—with uncertainties—allows Bayesian or Pareto planners to trade activity against stability, transport, safety, and manufacturability, advancing along frontiers that matter to devices and regulators (Yang *et al.*, 2024; Hörmann, Ambrosio, & Pasquarello, 2019).

Takeaway. A modern discovery stack is not data versus physics but data plus physics, architected across scales. Atomistic theory sets elementary rules; interfacial models define operating microenvironments; kinetics and transport predict operando behavior; and CG/continuum models carry insights to application scale. With uncertainty explicitly modeled, the pipeline proposes only those candidates whose physics permits high performance and real-world viability (Motagamwala & Dumesic, 2020).

12. Best Practices: Reporting & Reproducibility

Reproducible science and trustworthy deployment start with complete, standardized reporting. At minimum, every study should disclose data splits (how the training/validation/test partitions were constructed and why), random seeds and the number of independent runs, uncertainty quantification (UQ) and calibration metrics, baselines (well-tuned and appropriate to the task), and the compute budget and hardware used to obtain results. Reporting a single lucky run obscures variance due to initialization and nondeterminism; instead, authors should provide mean, standard deviation, and confidence intervals across multiple seeds, and pre-register the stopping criterion and early-stopping policy to avoid post-hoc cherry-picking (Henderson, Islam, Bachman, Pineau, Precup, & Meger, 2018; Dodge, Gururangan, Card, Schwartz, & Smith, 2019). UQ should move beyond raw accuracy to include calibration (e.g., expected calibration error) and

coverage under distribution shift, because well-calibrated probabilities are the substrate for active learning and risk-aware decision making in closed loops (Guo, Pleiss, Sun, & Weinberger, 2017). Finally, the compute and carbon cost of training and selection—datasets, model sizes, hyperparameter sweeps—should be summarized alongside accuracy; “Green AI” encourages comparisons at fixed compute budgets and transparency about energy sources to incentivize efficient methods (Schwartz, Dodge, Smith, & Etzioni, 2020).

Model cards and datasheets are now essential artefacts that travel with a trained model or dataset. A good model card states intended use, out-of-scope uses, training data characteristics, known biases, applicability domains, and failure modes, plus performance broken down by subgroups or regimes when relevant (Mitchell *et al.*, 2019). Datasheets for datasets document how data were collected, filtered, labeled, and split; legal and ethical constraints; and known hazards such as label ambiguity or class imbalance (Gebru *et al.*, 2018). For scientific ML in chemistry, materials, and bio, these ideas translate directly: disclose assay protocols or DFT settings that generated labels; cite software versions and pseudopotentials; list preprocessing (standardization, conformer generation, charge states); and include **assay** drift or batch effects if data were collected over time. Both artefacts should explicitly flag the applicability domain—for example, chemical space boundaries defined by Bemis–Murcko scaffolds or materials prototypes—and summarize OOD (out-of-distribution) performance where tested.

Open science is not just a philosophy but a set of release practices. Code and data should ship with explicit licenses—permissive when possible, for broad reuse—and, for datasets and models, archived with DOIs so versions are citable and immutable. Reproducible execution is dramatically easier when authors provide containers (e.g., Docker) or environment specifications pinned to exact versions, and when they include scripts to reproduce every table and figure end-to-end. For computational chemistry and materials, releases should include seed lists, hyperparameter grids, and, when applicable, trained weights for baselines so reviewers and downstream users can verify claims without rerunning expensive training (Boettiger, 2015; Pineau *et al.*, 2021). Public leaderboards can help track progress, but they must be curated to prevent overfitting to a static test set; hidden or rolling test sets, and periodic refreshes, reduce gaming while preserving comparability (Recht, Roelofs, Schmidt, & Shankar, 2019).

Evaluation design is where many otherwise solid studies fail. Data leakage—any flow of information from the test set into training or model selection—artificially inflates metrics. In molecular AI, random splits routinely let near-duplicates (or close analogs) leak across splits; scaffold splits mitigate this by holding out

chemical scaffolds never seen in training, offering a more realistic generalization test (Wu *et al.*, 2018). In materials discovery, structure-aware splits must ensure that compositions, prototypes, or even space groups in the test set are not trivial variants of training exemplars; failing that, reported performance mostly reflects interpolation. Across domains, authors should probe distribution shift explicitly—temporal splits, domain splits (e.g., assay lab or instrument), or geographic splits—and report degradation and calibration drift relative to i.i.d. conditions (Recht *et al.*, 2019). When hyperparameters are tuned, the validation set must be strictly disjoint from the test set, and any ablation or ensembling done after peaking at test results should be labeled post-hoc and re-validated.

Because many scientific targets are **noisy**—from assay variability to DFT functional errors—studies should quantify and propagate aleatoric and epistemic uncertainty. Techniques such as temperature scaling or Dirichlet calibration improve probability estimates; ensembles, Monte Carlo dropout, deep evidential regression, and conformal prediction provide well-defined coverage guarantees or uncertainty intervals that are actionable in downstream decision making (Guo *et al.*, 2017; Angelopoulos & Bates, 2022). Crucially, uncertainty reporting should include coverage vs. set size plots and risk-coverage curves, not just a single ECE number, and it should be repeated under the same distribution-shift settings used for accuracy.

To make results durable, every paper should include a Reproducibility Checklist covering data access, preprocessing, model specification, training regimen, hardware, runtime, and exact commands. Hyperparameters should be reported as complete config files (not prose), and early-stopping criteria and patience values should be specified. For stochastic training, authors should release the exact train/validation/test indices used (or the code and seed that deterministically generates them) so that later work can reproduce comparisons. When claims hinge on statistical significance, papers should report the test used, the effect size, and the number of trials; tiny but statistically significant gains with very large sample sizes should be contextualized against computational cost and complexity.

Benchmarking culture benefits from multiple, diverse baselines and from unit-consistent metrics. For example, in adsorption or catalysis tasks, include physically motivated baselines (simple descriptors, linear models) and report error units (e.g., eV, $\text{kJ}\cdot\text{mol}^{-1}$) alongside dimensionless metrics. In drug discovery, include docking or physics-based baselines where appropriate and stratify performance by chemotype; in remediation, stratify by matrix conditions. Negative results—failed transfer to a new assay, or degraded performance in a new water matrix—should be reported with the same care as positive results; they are often the

most valuable information to future practitioners. Finally, set deployment-minded thresholds: for any model intended to guide experiments, define acceptable confidence or coverage targets that would trigger escalation to higher-fidelity physics or wet-lab confirmation rather than blind action.

13. Case Studies

Catalysis — Single-atom catalyst discovered via active learning, validated by DFT, and scaled to an electrolyzer.

A carbon-supported single-atom catalyst (SAC) program for oxygen evolution illustrates the full closed loop from hypothesis to device. The team began with heteroatom-doped carbons hosting isolated M-N_x motifs and used an uncertainty-aware active-learning planner to batch-propose syntheses, guided by a calibrated graph model trained on prior *ex situ* spectra and *ab initio* adsorption descriptors. Each cycle fed standardized rotating-disk tests at matched electrolyte and pH, and top candidates advanced to density-functional theory with nudged-elastic-band checks to confirm barrier trends consistent with microkinetic optima rather than overbinding artefacts (Motagamwala & Dumesic, 2020). Within ten iterations, the loop converged on a narrow neighborhood of coordination and co-dopant patterns whose local strain and ligand field positioned intermediates on the OER volcano while suppressing site blockage. Crucially, down-selection then shifted from half-cell metrics to membrane-electrode assembly (MEA) targets, letting the planner trade intrinsic activity against ink rheology, ionomer ratio, and through-plane porosity. Several half-cell “winners” fell away once mass-transport and humidity sensitivities were penalized. The final MEA sustained target current density at low cell voltage for extended hours with minimal agglomeration, consistent with the SAC mechanism inferred from operando spectroscopy and *ab initio* signatures. The case highlights three lessons: (i) sequential learning accelerates exploration only when paired with faithful physics checks; (ii) device-level constraints must be embedded as objectives, not applied post hoc; and (iii) reporting uncertainty and variance across replicates is essential to avoid over-claiming incremental gains (Shields *et al.*, 2021; Motagamwala & Dumesic, 2020).

Drug — Generative lead optimized with alchemical free energies and ADMET filters to an *in vivo* efficacy signal.

A kinome-biased generator seeded chemotypes for an ATP-competitive kinase program, but proposals advanced only if they cleared synthesis-feasibility constraints and a multi-task predictor returned potency with calibrated uncertainty plus early-risk proxies (hERG, CYP, reactivity). Poses were generated and rescored; near-ties triggered alchemical free-energy refinement (FEP/TI) using ensemble protocols to resolve $\sim 1 \text{ kcal}\cdot\text{mol}^{-1}$ differences—large enough to reorder top-10 ranks (York *et al.*, 2023). Make-test-analyze

cycles—each logging negative results with equal fidelity—mapped the local SAR while uncertainty routing periodically forced exploration beyond the initial scaffold basin (Ekins *et al.*, 2019). In parallel, ADMET triage moved from *in-silico* to *in vitro* (microsomal stability, CYP panels, solubility/permeability, early cardiotoxicity). A potency-neutral scaffold edit that improved polarity reduced hERG risk and lifted metabolic stability, enabling exposures compatible with a short *in vivo* pharmacology study where one analog produced a statistically significant efficacy signal at tolerated doses. Retrospective attribution showed inflection points aligned with physics-based decisions—FEP/TI reversals of ML ties and uncertainty-driven sampling of underrepresented chemotypes—rather than brute-force enumeration. The case demonstrates that generative ideation must be coupled to calibrated prediction, targeted physics, and staged ADMET to turn *in silico* promise into *in vivo* signal quickly while managing compute and wet-lab risk (York *et al.*, 2023; Ekins *et al.*, 2019).

Environment — MOF/COF platform for PFAS capture with LCA vs. activated carbon and a regeneration plan.

Facing mixed per- and polyfluoroalkyl substances (PFAS), the program targeted water-stable Zr-MOFs (UiO-66 derivatives) and fluorophilic COFs tuned for head-group electrostatics and tail-framework affinity. Batch isotherms in realistic matrices (hardness, competing anions, natural organic matter) identified candidates with fast kinetics and high capacities; column tests quantified breakthrough bed volumes and regeneration via salt/solvent swings or electro-regeneration. UiO-66 variants delivered robust affinity for long-chain PFAS, while tailored COFs showed advantages for short-chain species due to pore architecture. Because PFAS management is prone to burden shifting, the team ran a life-cycle comparison at a common functional unit (liters treated to below limits): baseline granular activated carbon versus best-MOF and best-COF. Inventories covered synthesis precursors, solvent recovery, transport, pressure drop, regeneration chemicals/energy, and end-of-life. While activated carbon remained competitive on cost and embodied energy, the optimized MOF reduced media consumption and waste when regeneration and media lifetime were credited; COFs narrowed the gap for short-chain PFAS but required solvent-recovery improvements to retain advantage (Wang, DeWitt, Higgins, & Cousins, 2017; Dong, Tu, & Zheng, 2020). Sensitivity analysis flagged electricity mix and solvent recycle as dominant levers. The techno-environmental outcome was pragmatic: a hybrid train—GAC bulk removal followed by MOF/COF cartridges targeted to the local PFAS profile—minimized energy per log removal and landfill burden, provided operations tracked fluoride balance and screened for precursor-to-short-chain by-products to keep “destruction” claims auditable (Qiao, Guo, & Sun, 2023; Crini & Lichtfouse, 2019).

14. Grand Challenges & Outlook

The next decade will be defined less by algorithmic novelty than by our ability to learn from scarcity and to elevate negative results into first-class training signals. In chemistry, materials, and bio, data are sparse, noisy, and biased toward successes; failed syntheses, unstable formulations, and null bioassays are rarely curated with the same care as “wins,” yet they are disproportionately informative for calibrating risk and shaping acquisition functions in closed loops. Community norms must shift toward routine release of structured nulls, with datasheets that document assay drift, batch effects, and provenance, and with conformal or Bayesian machinery that converts uncertainty into actionable coverage guarantees (Gebru *et al.*, 2018; Pineau *et al.*, 2021; Angelopoulos & Bates, 2022). Without this, sample-efficient planners overfit to historical luck and under-explore the regimes where breakthroughs usually hide.

A second, persistent obstacle is building generalizable, uncertainty-aware models that cross domains and distribution shifts. Random splits and single-seed reporting are no longer defensible; scaffold-, temporally-, and site-aware splits should become the default in molecular, materials, and process datasets, accompanied by calibration metrics and risk-coverage curves rather than single accuracies (Wu *et al.*, 2018; Recht, Roelofs, Schmidt, & Shankar, 2019; Guo, Pleiss, Sun, & Weinberger, 2017). Practically, this means multi-task and transfer learners that ingest heterogeneous modalities—graphs, grids, spectra, sequence, operando streams—while emitting well-calibrated predictive intervals that tell planners when to escalate to physics or experiment. It also means guardrailed surrogates that encode symmetries, conservation laws, and correct far-field limits so extrapolation fails gracefully rather than confidently wrong (Shields *et al.*, 2021).

The frontier is no longer “accuracy only,” but multi-objective sustainability: performance and safety and cost. Discovery stacks must optimize activity/selectivity alongside PMI/E-factor, solvent guides, and fast LCA vectors so acquisition decisions reflect whole-system impact, not just a yield or ΔG (Sheldon, 2018; Hollberg & Ruth, 2021). For nano-enabled systems, eco-tox and corona dynamics should be encoded as constraints with OECD-aligned characterization plans, so risky shapes, coatings, or leaching behaviors are down-weighted early (Drasler *et al.*, 2017; Abdolapur Monikh *et al.*, 2020). The cultural shift is to treat sustainability numbers like any KPI: plotted, benchmarked, and traded on a Pareto front—rather than relegated to the supplement.

Trust and interpretability will determine adoption in the clinic, plant, and municipal utility. For regulators and operators, the question is not only “does it work?” but “why should we believe, it will keep working

here, now?” Model cards and datasheets must move into scientific ML with domain-specific content: assay conditions, DFT settings, pseudopotentials, instrument versions, and applicability domains by scaffold/prototype, plus OOD behavior under known shifts (Mitchell *et al.*, 2019; Gebru *et al.*, 2018). Interpretable mechanisms—substructure saliency, counterfactuals, mechanistic ablations, and microkinetic attribution—should be demanded when decisions carry safety or regulatory consequences. Above all, calibration matters more than raw AUC: a well-calibrated 0.80 AUC model with coverage guarantees is more deployable than a brittle 0.85 that hides its uncertainty (Guo *et al.*, 2017; Angelopoulos & Bates, 2022).

Scaling from high-throughput experimentation (HTE) to plant or clinic remains a chasm. Self-driving campaigns excel at local optimization under tight control; translation collapses when unit operations, mass/heat transfer, supply chains, and GxP data integrity enter. Best practice is to carry deployment constraints into the loop: rheology windows and ionomer ratios for MEAs; microfluidic mixing and critical quality attributes for LNPs; fouling, pressure drop, and regenerant handling for water trains. Optimization targets should be device-level (cell voltage at current density; cycle life to 20% fade; $\text{\$}\cdot\text{m}^{-3}$ treated) and paired with TEA/LCA so plant-scale viability evolves with the chemistry (Tom *et al.*, 2024; Hollberg & Ruth, 2021). Real-world validation needs multi-site pilots and prospective studies with pre-registered metrics and stop rules; “hero plots” from one lab are no longer enough.

On the compute horizon, quantum advantage and exascale HPC will be complementary rather than substitutive. Exascale enables many-query physics—GW/BSE, NEB ensembles, long-time molecular dynamics, and uncertainty propagation—so that active learners can afford the expensive labels that actually change decisions. Near-term quantum devices are unlikely to replace this, but they may carve out niche accelerants: strongly correlated fragments, compact ansätze for excited states, or discrete optimization within retrosynthesis and materials packing—provided error mitigation and problem encoding are robust (Tom *et al.*, 2024). The realistic outlook is hybrid: HPC-backed surrogates and ML interatomic potentials do the bulk lifting, with quantum subroutines called as specialized oracles where classical scaling is the bottleneck.

A final challenge is coordination: we need a community roadmap that makes progress legible, comparable, and cumulative. Three ingredients stand out. First, shared testbeds—modular photoreactors, membrane skids, electrochemical stacks, and biophysics benches—with standardized schemas, drivers, and QC protocols, so methods can be drop-in evaluated under identical conditions (SiLA-style). Second, open reference workflows that run end-to-end—from data ingestion to plots—captured in containers with DOIs, so

claims can be rerun years later on new hardware (Boettiger, 2015; Pineau *et al.*, 2021). Third, prize challenges that reward not only peak accuracy but also calibrated uncertainty, robustness under shift, compute/carbon efficiency, and sustainability trade-offs. The prize metric could be composite: weighted performance, calibrated coverage, TEA/LCA penalty, and a reproducibility score derived from the reporting checklist.

What, concretely, should teams do next? Treat uncertainty as currency: quantify it, report it, and spend it where it buys the most information. Elevate negative results and operational constraints to first-class citizens in optimization. Couple physics and data with explicit information contracts across scales. Make sustainability part of the acquisition function, not the after-action review. And align artifacts—model cards, datasheets, containers, DOIs—with the expectation that others will build on your work. If the field adopts these norms, closed-loop discovery can move from spectacular one-offs to reliable, auditable pipelines that deliver deployable catalysts, therapeutics, and remediation technologies at the pace society now demands (Sheldon, 2018; Mitchell *et al.*, 2019; Pineau *et al.*, 2021; Tom *et al.*, 2024).

REFERENCES

- Abdolapur Monikh, F., *et al.*, (2020). Are we generating reliable, realistic, and relevant data from nano-ecotoxicological studies? *Environmental Science: Nano*, 7(8), 2170–2186.
- Agha, A., *et al.*, (2023). Microfluidic-assisted nanoparticle synthesis and scale-up considerations. *Micromachines*, 14, 115.
- Ahmad, H., Sarwar, M. A., Khan, A. N., Alvi, A., Riaz, U., Khawaja, L., Mudassar, M. A., Ahmad, A., ILTAF, & Khan, W. Z. (2025, August 7). Unified intelligence: A comprehensive review of the synergy between data science, artificial intelligence, and machine learning in the age of big data. *Scholars Journal of Engineering and Technology*, 13(8), 585–617. <https://doi.org/10.36347/sjet.2025.v13i08.001>
- Alsalloum, G. A., *et al.*, (2024). Digital twins of biological systems: A narrative review. *Frontiers in Physiology*, 15, 11342927.
- Alshawwa, S. Z., Abusharkh, S. E., Almashhadani, H. A., Alsaedi, A., Al-Brahim, H., & Almahallawi, A. M. (2022). Nanocarrier drug delivery systems: Characterization, applications, and challenges. *Polymers*, 14(8), 1607.
- Altae-Tran, H., Ramsundar, B., Pappu, A. S., & Pande, V. (2017). Low-data drug discovery with one-shot learning. *ACS Central Science*, 3(4), 283–293.
- An, G. (2022). Drug development digital twins for discovery, testing, and repurposing. *Frontiers in Systems Biology*, 2, 928387.

- Anekwe, I. M. S., Li, T., Salako, F. I., & Zhang, W. (2025). Unlocking catalytic longevity: A critical review of catalyst deactivation pathways and regeneration technologies. *Environmental Science: Advances*, 4, 100–130.
- Angelopoulos, A. N., & Bates, S. (2022). A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv Preprint*, arXiv:2107.07511.
- Angelopoulos, A. N., & Bates, S. (2023). A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *Foundations and Trends in Machine Learning*, 16(4), 494–591.
- Anselmo, A. C., & Mitragotri, S. (2019). Nanoparticles in the clinic: An update. *Bioengineering & Translational Medicine*, 4(3), e10143.
- Bartók, A. P., Kondor, R., & Csányi, G. (2013). On representing chemical environments. *Physical Review B*, 87(18), 184115.
- Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J., Kornbluth, M., ... & Kozinsky, B. (2022). E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13, 2453
- Baz, A., Comas-Vives, A., & López, N. (2021). Microkinetic modeling in electrocatalysis. *Journal of Catalysis*, 404, 744–763.
- Benison, K. C., *et al.*, (2022). A practical approach to evaluating solvent sustainability in pharmaceutical process development. *Organic Process Research & Development*, 26(9), 2148–2160.
- Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: Fair enough? *European Journal of Human Genetics*, 26, 931–936.
- Boettiger, C. (2015). An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*, 49(1), 71–79.
- Brown, N., Fiscato, M., Segler, M. H. S., & Vaucher, A. (2019). GuacaMol: Benchmarking models for de novo molecular design. *Journal of Chemical Information and Modeling*, 59(3), 1096–1108.
- Caruana, R. (1997). Multitask learning. *Machine Learning*, 28(1), 41–75.
- Chanussot, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., ... & Ulissi, Z. W.
- Chen, J., Wang, Y., Li, X., & Chen, G. (2024). Selective and stable CO₂ electroreduction at high rates via polymer-regulated water management. *Nature Communications*, 15, 50269.
- Chen, Y., Ji, S., Zhao, S., Chen, W., Dong, J., Cheong, W.-C., ... Li, Y. (2021). Single-atom catalysts: Synthetic strategies and electrochemical applications. *Joule*, 5(4), 1044–1065.
- Cheng, Y.-L., Luo, J., & Cheng, J. (2022). Examination of the Brønsted–Evans–Polanyi relationship at constant potential for electrocatalysis. *Physical Chemistry Chemical Physics*, 24, 12419–12427.
- Choudhary, K., & Tavazza, F. (2020). JARVIS-DFT: An integrated infrastructure for density functional theory calculations. *Scientific Data*, 7, 227.
- Christensen, M., *et al.*, (2021). Data-science-driven autonomous process optimization. *Communications Chemistry*, 4, 112.
- Coley, C. W., Eyke, N. S., & Jensen, K. F. (2019). Autonomous discovery in the chemical sciences part I: Progress. *Angewandte Chemie International Edition*, 59(51), 22858–22893.
- Coley, C. W., Green, W. H., & Jensen, K. F. (2018). SCScore: Synthetic complexity learned from a reaction corpus. *Journal of Chemical Information and Modeling*, 58(2), 252–261
- Coley, C. W., Jin, W., Rogers, L., Jamison, T. F., & Jaakkola, T. (2019). A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical Science*, 10, 370–377.
- Coley, C. W., Green, W. H., & Jensen, K. F. (2021). RDChiral and the Open Reaction Database: Toward standardization of chemical reaction data. *Journal of Chemical Information and Modeling*, 61(10), 5417–5426.
- Crini, G., & Lichtfouse, E. (2019). Advantages and disadvantages of techniques used for wastewater treatment. *Environmental Chemistry Letters*, 17, 145–155.
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197.
- Diercks, C. S., & Yaghi, O. M. (2017). The atom, the molecule, and the covalent organic framework. *Science*, 355(6328), eaal1585.
- Ding, L., Wei, Y., Li, L., Zhang, T., Wang, H., Xue, J., ... Jiang, L. (2020). MXene molecular sieving membranes for highly efficient gas separation. *Nature Communications*, 11, 3247.
- Dodge, J., Gururangan, S., Card, D., Schwartz, R., & Smith, N. A. (2019). Show your work: Improved reporting of experimental results. In *Proceedings of the EMNLP Workshop on Reproducibility*.
- Dong, Y., Tu, B., & Zheng, H. (2020). Metal–organic frameworks for water remediation: From adsorption to catalysis. *Chemical Engineering Journal*, 399, 125791.
- Dunn, A., Wang, Q., Ganose, A., Dopp, D., & Jain, A. (2020). Benchmarking materials property prediction methods: Matbench test set and Automatminer reference algorithm. *npj Computational Materials*, 6, 138.
- Durumeric, A. E., Vani, B. P., & Onuchic, J. N. (2023). Machine-learned coarse-grained models for

- soft matter and biomolecules. *Annual Review of Physical Chemistry*, 74, 313–338. European Chemicals Agency. (2020). *Nanomaterials under REACH and CLP: Technical and regulatory guidance*.
- Ekins, S., Puhl, A. C., Zorn, K. M., Lane, T. R., Russo, D. P., Klein, J. J., ... Karanicolas, J. (2019). Exploiting machine learning for end-to-end drug discovery and development. *Clinical Pharmacology & Therapeutics*, 106(1), 89–98.
 - Ertl, P., & Schuffenhauer, A. (2009). Estimation of synthetic accessibility score of drug-like molecules. *Journal of Cheminformatics*, 1, 8.
 - EU Publications Office. (2017). *Technology readiness levels (TRL)*.
 - Finkbeiner, M., et al., (2020). Life cycle sustainability assessment—What is it and what are its challenges? *Sustainability*, 12(20), 861
 - Forzatti, P. (1999). Catalyst deactivation. *Catalysis Today*, 52(2–3), 165–181.
 - Frazier, P. I. (2018). A tutorial on Bayesian optimization. *arXiv Preprint*, arXiv:1807.02811.
 - Frišić, T., Mottillo, C., & Titi, H. M. (2020). Mechanochemistry for synthesis. *Angewandte Chemie International Edition*, 59(3), 1018–1029.
 - Frost, J. M. (2017). Calculating polaron mobility in halide perovskites. *Physical Review B*, 96(19), 195202.
 - Garcia-Segura, S., Ocon, J. D., & Chong, M. N. (2018). Electrochemical advanced oxidation processes (EAOPs) for persistent organic pollutants: A review. *Journal of Chemical Technology & Biotechnology*, 93, 309–323.
 - Garrido, A., Lepailleur, A., Mignani, S. M., & Zagury, J.-F. (2020). hERG toxicity assessment: Useful guidelines for drug design. *European Journal of Medicinal Chemistry*, 195, 112290.
 - Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2018). Datasheets for datasets. *Communications of the ACM*, 64(12), 1–17.
 - Ghosh, S., Zhou, J., & Wong, B. M. (2020). Strain engineering at the nanoscale: Electronic structure and catalysis. *ACS Nano*, 14(11), 14676–14685.
 - Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., & Dahl, G. E. (2017). Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 1263–1272).
 - Göltl, F., et al., (2022). Generalized Brønsted–Evans–Polanyi relationships for heterogeneous catalysis. *ChemCatChem*, 14, e202201108.
 - Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., ... Aspuru-Guzik, A. (2018). Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2), 268–276.
 - Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 1321–1330).
 - Häse, F., Aldeghi, M., Hickman, R. J., Roch, L. M., & Aspuru-Guzik, A. (2020). Gryffin: Bayesian optimization of categorical variables with expert knowledge. *arXiv Preprint*, arXiv:2003.12127.
 - Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., & Meger, D. (2018). Deep reinforcement learning that matters. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 3207–3214).
 - Henkelman, G., Uberuaga, B. P., & Jónsson, H. (2000). A climbing-image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113(22), 9901–9904.
 - Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In *Proceedings of the 34th Conference on Neural Information Processing Systems*.
 - Hoogeboom, E., Satorras, V. G., Vignac, C., & Welling, M. (2022). Equivariant diffusion for molecule generation. In *Proceedings of the 39th International Conference on Machine Learning*.
 - Hörmann, N. G., Ambrosio, F., & Pasquarello, A. (2019). Absolute band alignment at semiconductor–water interfaces. *npj Computational Materials*, 5, 100.
 - Hollberg, A., & Ruth, J. (2021). Visualization in life cycle assessment results: A review and recommendations. *Journal of Industrial Ecology*, 25(6), 1437–1452.
 - Huang, K., Fu, T., Gao, W., Zhao, Y., Roohani, Y., Leskovec, J., ... Zitnik, M. (2021). Therapeutics Data Commons: Machine learning datasets and tasks for therapeutics. *Nature Machine Intelligence*, 4, 252–264.
 - Huang, Y., He, Z., & Zhong, Y. (2019). Sulfate radical-based advanced oxidation processes for organic wastewater treatment: A review. *Chemical Engineering Journal*, 372, 836–851.
 - Huo, H., & Rupp, M. (2017). Unified representation of molecules and crystals for machine learning. *arXiv Preprint*, arXiv:1704.06439.
 - Jiang, T., Sun, K., Shaydulin, R., Lubasch, M., & Liu, J.-G. (2024). Error mitigation in variational quantum algorithms. *Physical Review Research*, 6, 033069.
 - Jin, W., Barzilay, R., & Jaakkola, T. (2018). Junction tree variational autoencoder for molecular graph generation. In *Proceedings of the 35th International Conference on Machine Learning* (pp. 2323–2332).
 - Kearnes, S. M., et al., (2021). The Open Reaction Database. *Journal of the American Chemical Society*, 143(45), 18820–18826.

- Kearns, P., *et al.*, (2021). Governance frameworks for nanotechnology: Fostering responsible innovation. *Nature Nanotechnology*, 16(8), 800–810.
- Kekessie, S. E., *et al.*, (2024). Quantifying solvent impacts with process mass intensity: Case studies from route scouting to launch. *Green Chemistry*, 26(3), 1452–1467.
- Khan, S. A., Günther, P. M., & Jensen, K. F. (2021). Microfluidic synthesis of nanomaterials. *Chemical Reviews*, 121(22), 13230–13301.
- Klicpera, J., Groß, J., & Günnemann, S. (2020). Directional message passing for molecular graphs (DimeNet). In *Proceedings of the International Conference on Learning Representations*.
- Krenn, M., Häse, F., Nigam, A., Friederich, P., & Aspuru-Guzik, A. (2020). SELFIES: A 100% robust molecular string representation. *Machine Learning: Science and Technology*, 1(4), 045024.
- Kulkarni, J. A., Witzigmann, D., Thomson, S. B., Chen, S., Leavitt, B. R., Cullis, P. R., & van der Meel, R. (2021). The current landscape of lipid nanoparticles for mRNA delivery. *Nature Nanotechnology*, 16(6), 630–643.
- Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In *Proceedings of the 31st Conference on Neural Information Processing Systems* (pp. 6402–6413).
- Leonzio, G. (2024). CO₂ electrochemical reduction: A state-of-the-art review of figures of merit and device engineering. *Journal of CO₂ Utilization*, 85, 102263.
- Li, J., & Zeng, Q. (2019). Colloidal quantum dots: From synthesis to applications. *Accounts of Chemical Research*, 52(9), 2406–2416.
- Li, P., Zeng, Z., Wang, B., & Chen, Y. (2018). Enzyme–nanoparticle biocatalysts: From synthesis to applications. *Chemical Society Reviews*, 47(22), 7815–7851.
- Liu, X., Testa, A., & Huang, Y. Y. S. (2021). Considerations for the scale-up manufacture of nanotherapeutics: A critical step for technology transfer. *View*, 2(6), 20200190.
- Liu, X., Zhang, Y., & Lowry, G. V. (2023). Eco-corona formation and implications for environmental risks of nanomaterials. *Environmental Science & Technology*, 57(2), 810–824.
- Liu, Z., Li, Y., Han, L., Li, J., Liu, J., Zhao, Z., ... Wang, R. (2017). PDB-wide collection of binding data: Current status of the PDBbind database. *Bioinformatics*, 31(3), 405–412.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st Conference on Neural Information Processing Systems* (pp. 4765–4774).
- Mardirossian, N., & Head-Gordon, M. (2017). Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Molecular Physics*, 115(19), 2315–2372.
- McCloskey, K., Taly, A., Monti, F., & Brenner, M. P. (2019). Using attribution to decode binding mechanism in neural network models for chemistry. *Proceedings of the National Academy of Sciences*, 116(24), 11624–11629.
- Meggiolaro, D., & De Angelis, F. (2020). Polarons in lead-halide perovskites. *ACS Energy Letters*, 5(5), 1556–1566.
- Mehta, M., Chen, J., & White, M. A. (2023). Lipid-based nanoparticles for drug/gene delivery. *ACS Materials Au*, 3(3), 238–260.
- Melander, M. M., Sundararaman, R., & Groß, A. (2024). Constant-inner-potential density-functional theory for electrochemical interfaces. *npj Computational Materials*, 10, 112.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229).
- Motagawala, A. H., & Dumesic, J. A. (2020). Microkinetic modeling: A tool for rational catalyst design. *Chemical Reviews*, 120(23), 12379–12443.
- Nizam, N. U. M., Arshad, A., & Sultana, S. (2021). A content review of life cycle assessment of nanomaterials. *Nanomaterials*, 11(12), 3324.
- NORMAN Network. (2021). *SusDat: The NORMAN Suspect List Exchange* [Dataset note; latest major update 2021].
- Papamakarios, G., Pavlakou, T., & Murray, I. (2017). Masked autoregressive flow for density estimation. In *Proceedings of the 31st Conference on Neural Information Processing Systems* (pp. 2338–2347).
- Park, C. W., Kornbluth, M., Wood, B., & Jain, A. (2023). Mat2Vec and beyond: Learned embeddings for materials science. *Annual Review of Materials Research*, 53, 541–563.
- Paudel, H. P., *et al.*, (2022). Quantum computing and simulations for energy applications. *ACS Engineering Au*, 2(4), 274–287.
- Pérez-Reyes, A., Téllez, C., & Coronas, J. (2021). Mixed matrix membranes based on MOFs and COFs for gas separation. *Chemical Society Reviews*, 50(14), 7949–7990.
- Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., & Battaglia, P. (2021). Learning mesh-based simulation with graph networks. In *Proceedings of the International Conference on Learning Representations*.
- Pham, P., *et al.*, (2025). A comprehensive review of catalyst deactivation and regeneration. *Catalysis Today*, 438, 1–25.
- Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivière, V., Beygelzimer, A., d’Alché-Buc, F., ... Hutter, F. (2021).

- Improving reproducibility in machine learning research: A report from the NeurIPS 2019 reproducibility program. *Journal of Machine Learning Research*, 22, 1–20.
- Pyzer-Knapp, E. O., *et al.*, (2025). Foundation models for materials discovery—Current state and future directions. *npj Computational Materials*, 11, 61.
- Qiao, Z., Guo, Y., & Sun, W. (2023). Electrochemical destruction of per- and polyfluoroalkyl substances: Progress and perspectives. *Chemosphere*, 330, 138676.
- Qiu, Y., Zhang, Q., & Zhao, H. (2019). Membrane technologies for heavy metal and microplastics removal: Progress and challenges. *Journal of Membrane Science*, 586, 53–67.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks. *Journal of Computational Physics*, 378, 686–707.
- Recht, B., Roelofs, R., Schmidt, L., & Shankar, V. (2019). Do ImageNet classifiers generalize to ImageNet? In *Proceedings of the 36th International Conference on Machine Learning* (pp. 5389–5400).
- Rizzi, A., Rehbein, P., Zeller, F., & Hummer, G. (2021). Accurate free energies from mapped reference potentials. *The Journal of Physical Chemistry Letters*, 12(39), 9449–9455.
- Ruan, Y., *et al.*, (2022). AROPS: Automated reaction optimization with parallelized scheduling. *ChemRxiv* (preprint).
- Saal, J. E., Kirklin, S., Aykol, M., Meredig, B., & Wolverton, C. (2013). Materials design and discovery with high-throughput DFT: The Open Quantum Materials Database (OQMD). *JOM*, 65(11), 1501–1509.
- Schölkopf, B., Locatello, S., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., & Bengio, Y. (2021). Toward causal representation learning. *Proceedings of the IEEE*, 109(5), 612–634.
- Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A., & Müller, K.-R. (2018). SchNet: A deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24), 241722.
- Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63.
- Segler, M. H. S., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*, 555, 604–610.
- Settles, B. (2009). Active learning literature survey. *University of Wisconsin–Madison Technical Report*.
- Sheldon, R. A. (2017). Metric of green chemistry and sustainability: Past, present, and future. *ACS Sustainable Chemistry & Engineering*, 5(7), 5527–5538.
- Sheldon, R. A. (2018). E-factor 25 years on: The rise of green chemistry and sustainability. *Green Chemistry*, 20(1), 18–43.
- Shields, B. J., *et al.*, (2021). Bayesian reaction optimization as a tool for chemical synthesis. *Nature*, 590(7844), 89–96.
- SiLA Consortium. (2019). SiLA 2: The emerging standard for lab automation. *Wiley Analytical Science* (feature article).
- Su, M., Yang, Q., Du, Y., Feng, G., Liu, Z., Li, Y., & Wang, R. (2020). Comparative assessment of scoring functions on updated benchmarks (CASF-2016/2020). *Journal of Chemical Information and Modeling*, 60(11), 5936–5953.
- Togo, A. (2023). First-principles phonon calculations with phonopy and phono3py. *Journal of the Physical Society of Japan*, 92(1), 012001.
- Tom, G., *et al.*, (2024). Self-driving laboratories for chemistry and materials science. *Chemical Reviews*, 124(13), 7137–7210.
- Tran, H. N., Ok, Y. S., & Sik, O. (2017). Activated carbon and biochar for water treatment: A critical review of adsorption mechanisms. *Bioresource Technology*, 246, 34–49.
- Tran, R., Ulissi, Z. W., & Chanussot, L. (2023). Progress and opportunities in machine learning for heterogeneous catalysis. *Chemical Reviews*, 123, 1514–1556.
- Voiry, D., Shin, H. S., Loh, K. P., & Chhowalla, M. (2018). Low-dimensional catalysts. *Nature Reviews Chemistry*, 2, 0105.
- Waller, P. J., Gándara, F., & Yaghi, O. M. (2019). Chemistry of covalent organic frameworks. *Accounts of Chemical Research*, 52(12), 3168–3178.
- Wang, W., & Wang, S. (2020). Graphitic carbon nitride (g-C₃N₄)-based photocatalysts for water purification under visible light. *Applied Catalysis B: Environmental*, 276, 119–123.
- Wang, Z., DeWitt, J. C., Higgins, C. P., & Cousins, I. T. (2017). A never-ending story of per- and polyfluoroalkyl substances (PFASs)? *Environmental Science & Technology*, 51(5), 2508–2518.
- Weidman, J. D., *et al.*, (2024). Quantum computing and chemistry. *Cell Reports Physical Science*, 5, 101624.
- Weiser, B. P., *et al.*, (2023). Machine-learning-augmented docking for cytochrome P450 inhibition prediction. *Digital Discovery*, 2(5), 1377–1391.
- 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, 160018.
- Wu, J., & Yang, H. (2020). Single-atom heterogeneous catalysts: Recent advances and challenges. *ACS Catalysis*, 10(14), 7798–7832.
- Wu, Z., Ramsundar, B., Feinberg, E. N., Gomes, J., Geniesse, C., Pappu, A. S., ... Pande, V. (2018). MoleculeNet: A benchmark for molecular machine learning. *Chemical Science*, 9(2), 513–530.

- 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.
- Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., ... Jaakkola, T. (2019). Are learned molecular representations ready for prime time? *Chemical Science*, 10(34), 8154–8160.
- Yang, T.-T., et al., (2024). The Bell–Evans–Polanyi relation for hydrogen evolution derived from electrochemical principles. *npj Computational Materials*, 10, 194.
- Ye, Z., & Zhang, P. (2021). Coarse-grained molecular simulations for soft materials: Recent advances. *Macromolecules*, 54(24), 11105–11125.
- Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D., ... Liu, T.-Y. (2021). Do transformers really perform badly on graphs? (Graphormer). In *Proceedings of the 35th Conference on Neural Information Processing Systems*.
- Ying, R., Bourgeois, D., You, J., Zitnik, M., & Leskovec, J. (2019). GNNExplainer: Generating explanations for GNNs. In *Proceedings of the 33rd Conference on Neural Information Processing Systems* (pp. 9244–9255).
- York, D. M., Gallicchio, E., Mobley, D. L., ... Shirts, M. R. (2023). Modern alchemical free-energy methods for drug discovery. *ACS Physical Chemistry Au*, 3(5), 902–921.
- Zhang, C., Chen, Q., & Quigg, A. (2024). Trophic transfer and transformation of engineered nanomaterials in aquatic food webs. *TrAC Trends in Analytical Chemistry*, 170, 117061.
- Zhao, W., Tan, C., Zhang, H., & Lau, S. P. (2020). Emerging ultrathin two-dimensional nanomaterials for electrocatalysis. *Small Methods*, 4(6), 1900586.
- Zhou, Z., Kearnes, S., Li, L., Zare, R. N., & Riley, P. (2019). Optimization of molecules via deep reinforcement learning. *Scientific Reports*, 9, 10752.
- Hussain, S., Yasmin, R., Chouhdary, A. N., ... Khan, W. Z. (2025, January). Atomic nuclei to nanostructures: Harnessing the convergence of nuclear physics, particle dynamics, and nanotechnology to transform energy, revolutionize medicine, and advance environmental sustainability. *GSAR Journal of Multidisciplinary Studies*, 4(1), 87–117.
- Khan, W. Z. (2025, May). *Targeted delivery and responsive release of protein drugs to the brain using GSH-responsive SiO₂ nanoparticles* (Unpublished master's thesis). University of Education, Lahore.
- Khan, W. Z. (2025, October). *Role of spin-orbit coupling in Bi-rich to In-rich Cs₂Ag(In_xBi_{1-x})Cl₆: Electronic and optical design by DFT*. Computational Materials Science. Advance online publication.
- Khan, W. Z., Butt, G., & Altaf, F. (2017). Efficacy and safety of intense pulsed light in the treatment of mild-to-moderate acne vulgaris. *Journal of Surgical Dermatology*, 2(T1), *Lasers in Dermatology* (themed issue).
- Khan, W. Z., Fatima, M., Ahmad, S., ... Tabassam, M. N. (2025, January). Advancing diagnostic medical physics: AI, MR-guided radiotherapy, and 3D printing in modern healthcare. *Dialogue SSR – Science, Society and Research*, 3(1). (Indexed in the journal's 2025 volume listing.)
- Khan, W. Z., Haseeb, M., Khan, S. N., ... Ahmad, S. (2025, April). Quantum materials: Key to advancing quantum computing by enhancing stability, scalability, and error resistance through superconductors, topological insulators, and 2D materials for scalable systems. Manuscript in preparation / preprint. (Public listing without final journal details.)
- Khan, W. Z., Yasmin, R., Akbar, H., ... Anwar, A. (2025, February). Laser technology: Bridging historical milestones and modern applications in science, industry, and sustainability. *Global Scientific and Academic Research Journal of Multidisciplinary Studies*, 4(1), 70–84. <https://doi.org/10.5281/zenodo.14770605>
- Leghari, H. M., Shah, I. H., Shahzad, F., Ul Haq, I., Mudassar, M. A., Ashfaq, F., Ghazal, H., Munawar, A., Shehzadi, S., & Khan, W. Z. (2025, August 7). Enhancing perovskite solar cell efficiency via tunable Ag nanoparticle-integrated SnO₂ transport layers: Mechanistic insights and air-fabrication approach. *Scholars Journal of Physics, Mathematics and Statistics*, 12(7), 254–267. <https://doi.org/10.36347/sjpms.2025.v12i07.002>
- Malik, A., Liaqat, M., ur Rahman, M., ... Khan, W. Z. (2025, February 23). Nanotechnology for perovskite solar cells: Solving efficiency, stability, and energy storage challenges. Article / preprint. (Public PDF available; journal not clearly specified in the record.)
- Mudassar, M. A., Khalid, A., Ashraf, A., Barkat, U., Aslam, M. A., Talha, H. M., Abdullah, M., Munawar, A., Ghazal, H., & Khan, W. Z. (2025, July 18). Theoretical design and optoelectronic analysis of lead-free CsPbX₃/Cs₂SnX₆ core-shell perovskite nanocrystals for enhanced stability and charge dynamics. *Scholars Journal of Physics, Mathematics and Statistics*, 12(6), 226–239. <https://doi.org/10.36347/sjpms.2025.v12i06.004>
- Nazar, M., Tabassam, M. N., Irfan, A., Liaqat, M., Ameen, M., Chaudhary, H. M., Murad, K., Alam, S., Sohaib, M., & Khan, W. Z. (2025, April 8). DFT study of optoelectronic and thermoelectric properties of halide double perovskite Rb₂TlSbX₆ (X = Cl, Br, I) for solar cell applications. *Scholars Journal of Engineering and Technology*, 13(4), 208–222. <https://doi.org/10.36347/sjet.2025.v13i04.002>
- Rafiq, M. N., Liaqat, M., Sadiqa, M., ... Khan, W. Z. (2025, April). Enhanced dielectric and piezoelectric properties of BaTiO₃-infused K_{0.3}Na_{0.2}Bi_{0.5}TiO₃ ceramics for high-frequency

applications. Manuscript / preprint. (Journal metadata not publicly verifiable.)

- Shafique, A. Y., Nazar, M., Liaqat, M., ... Khan, W. Z. (2025, February). Synthesis and characterization of La-doped BNBT lead-free ceramics with enhanced piezoelectric and dielectric properties. Research article (preprint). (Journal metadata not publicly verifiable.)
- Shakeel, R., Ahmad, A., Mushtaq, H., ... Khan, W. Z. (2025, August). Multifunctional nanomaterials in materials chemistry: A cross-disciplinary review of applications in energy, environment, catalysis, and biomedicine. *Scholars Journal of Engineering and Technology*, 13(8), 673–707. <https://doi.org/10.36347/sjet.2025.v13i08.006>
- Sharjeel, M., Tariq, S., Ahmad, M., Aslam, A., Iltaf, Khawaja, L., ... Khan, W. Z. (2025, September 30). An interdisciplinary review of modern computer science: Trends and challenges in AI, cybersecurity, cloud, blockchain, IoT, data science, NLP, vision, software engineering, and quantum computing. *Scholars Journal of Engineering and Technology*, 13(9), 768–793. <https://doi.org/10.36347/sjet.2025.v13i09.005>
- U.S. Environmental Protection Agency. (2024). *TSCA actions on nanoscale materials and PFAS: Program overview*.
- U.S. Food and Drug Administration. (2022). *Considerations for the use of nanomaterials in drug products*.
- Zahid, T., Shakeel, R., Asif, H. M., ... Khan, W. Z. (2025, July). Brain-targeted delivery and redox-responsive release of protein therapeutics via GSH-sensitive SiO₂ nanoparticles. *Advanced Materials Research and Review*. Advance online publication. (Article page shows journal as *Advanced Materials Research and Review*.)
- Zubair, M. T. B., Khan, W. Z., Abbas, H., ... Safeer-ul-Islam. (2025, September). First demonstration of lead-free double perovskite Cs₂NaBiI₆ as a dual-function electrode for high-performance, long-life photo-rechargeable lithium-ion batteries with superior solar-to-power conversion. *Scholars Journal of Physics, Mathematics and Statistics*, 12(8), 346–356. <https://doi.org/10.36347/sjpms.2025.v12i08.002>