
Entalpic’s White Paper: Accelerating Materials Discovery with AI

Entalpic Team

info@entalpic.ai

Materials underpin every modern technology, from the silicon in our laptops chips to catalysts in clean energy. Yet, discovering new materials remains slow, and resource-intensive. Entalpic is redefining this paradigm. Our AI-driven platform integrates predictive modeling, generative design, large language models (LLMs), active learning, automated quantum simulation workflows and experimental validation to enable faster, more targeted innovation in materials science. Specifically, it is designed to automate and optimise each stage of the discovery process, from hypothesis generation to experimental testing - reducing time-to-discovery by orders of magnitude in benchmark cases such as catalysts or battery cathodes. The synergy between computational and experimental methodologies ensures that our discoveries are not only theoretically promising, but also viable for real-world deployment, contributing to sustainable applications and high-performance materials development. In this white paper, we outline Entalpic’s scientific and technological approach, demonstrating how we empower chemists and industry leaders to accelerate R&D and unlock novel materials solutions.

Date: September 5, 2025

ENTALPIC

1 Introduction

From the Bronze Age to the Silicon Age, history has been shaped by materials breakthroughs. Today, materials remain central to the fabric of modern society, shaping the infrastructures we build, the energy we use, and the technologies we rely on. However, accelerating the transition to sustainable technologies requires a radical shift in how we discover new materials.

Traditional discovery methods, rooted in trial-and-error laboratory experiments guided by human intuition, are inherently very slow and resource-intensive. Even for the most skilled chemists, exploring the quasi-infinite-dimensional chemical space, with multiple constraints and properties to satisfy, is far beyond reach. This creates a major bottleneck in the development of materials and technologies critical to decarbonization [de Pablo et al., 2019, Chen et al., 2024]. Some quantum computational methods, like Density Functional Theory (DFT) [Kohn et al., 1996], exist and are accurate enough to be useful, but they are too resource-intensive for large-scale screening and precise chemical modelling.

This is why Machine learning (ML) now offers a leap forward, enabling DFT-level insights at much faster speeds [Zitnick et al., 2020]. Beyond making the evaluation of billions of materials candidates possible, ML models enable an intelligent and autonomous exploration of the chemical space, uncovering novel materials beyond human intuition and expanding the frontiers of materials discovery. Projects like The Materials Project [Jain et al., 2013], Open Catalysts Project [Chanussot et al., 2021], MatterGen [Zeni et al., 2023]

and GNoME [Merchant et al., 2023] are good examples in this new area, providing large-scale datasets and workflows to train predictive AI.

To address these challenges, Entalpic is building a next-generation AI-driven platform for materials discovery. Our pipeline is designed as a comprehensive automated discovery engine, integrating multimodal high-quality datasets, predictive and generative modelling, active learning loops, and experimental validation. Our experimental footprint is central to our approach, enabling exploration of materials candidates with a focus on practical and real-world applicability.

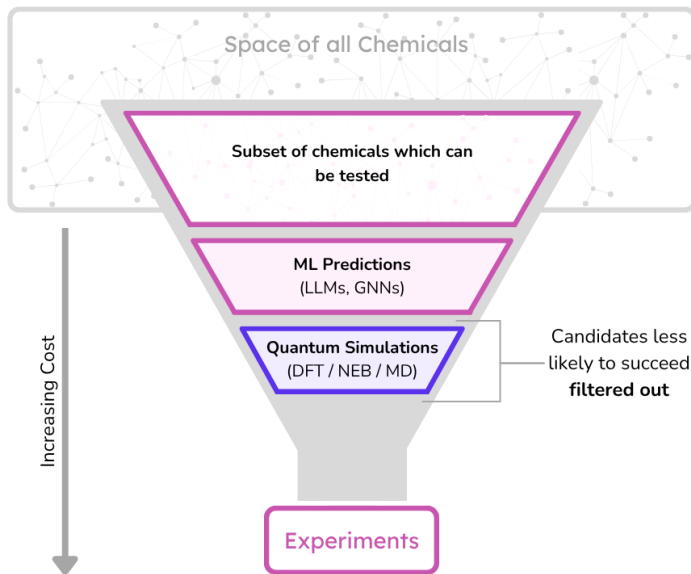


Figure 1 Simplified Schematic of Entalpic’s AI-driven screening funnel for material discovery. The funnel narrows the space of all possible chemicals to a small set of high-potential candidates by progressively applying computational filters. ML models such as graph neural networks (GNNs) and LLMs first screen vast chemical spaces. Promising candidates are then refined using higher-fidelity quantum simulations (e.g., DFT, Molecular Dynamics), after which only the most viable structures proceed to physical experiments. This multi-stage funnel reduces computational and experimental cost by eliminating low-potential candidates early in the pipeline.

2 Entalpic’s Platform Overview

Our AI-native platform connects five technological pillars:

1. **Predictive ML models** trained on multi-modal data for fast prediction of materials property.
2. **Generative models** capable of searching the chemical space efficiently and proposing stable, high-performance candidates.
3. **LLMs** that parse scientific literature and patents to extract domain knowledge and steer discovery.
4. **Experimental validation pipelines** for rapid synthesis, characterization, and performance testing, to bridge the gap between the computational and experimental worlds.
5. An **active learning loop, which brings these elements together**, incorporates feedback from both experiments and quantum simulations to refine our ML models iteratively. This continuously improves the quality and relevance of the proposed materials while reducing the time and cost of discovery.

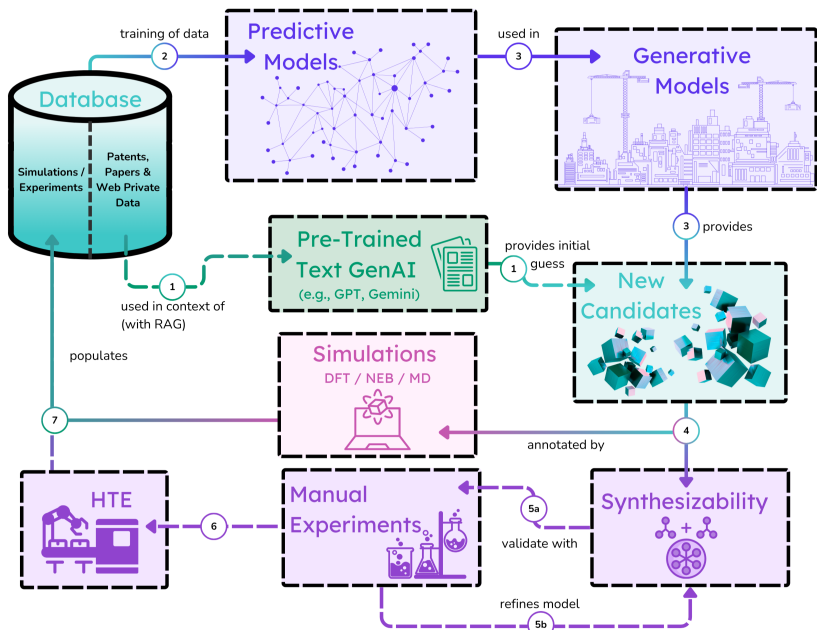


Figure 2 Overview of Entalpic’s AI-driven materials discovery platform — Our approach integrates predictive ML models trained on multi-modal data for fast prediction of material properties, generative models capable of searching the chemical space efficiently and proposing stable, high-performance candidates, an active learning loop that incorporates feedback from experiments and simulations to refine our models iteratively, and integration with experimental labs to bridge the gap between computational and experimental worlds.

2.1 Predictive Models : Screening of Materials Properties

Discovering new materials begins with the ability to quickly and accurately evaluate materials structures. Traditional laboratory experiments and physics-based simulations like DFT can do this first evaluation phase, but are too slow and computationally expensive to scale across the vastness of the chemical space.

To overcome this bottleneck, we **develop and deploy geometric Graph Neural Networks (GNNs)** [Kipf and Welling, 2016], a class of machine learning interatomic potentials that approximate quantum properties —such as formation energy, band gap, and adsorption energy —with near-DFT accuracy and orders-of-magnitude faster inference. These models represent atomic systems as 3D graphs, encoding both spatial geometry and chemical information to capture complex interatomic interactions [Duval et al., 2023].

We train those models on large datasets from [LeMaterial](#) —an open source effort we started together with [HuggingFace](#) —such as **LeMat-Bulk** [Siron et al., 2025], which standardizes, deduplicates, and harmonizes large-scale datasets, enabling reliable model training. By open-sourcing it, we aim to foster transparency, reproducibility, and collaboration in the materials science community.

At Entalpic, these predictive models form the backbone of our discovery engine, enabling high-throughput screening, candidate prioritization, and model-guided generative design. We continue to advance this foundation in three directions:

- **Data quality and scale** by pushing further our LeMaterial initiative every day by including diverse datasets, models and benchmarks with new partners.
- **Multimodal and multi-fidelity learning.** Our models integrate atomic graphs, crystallographic metadata and textual data, and learn from labels of varying fidelity (e.g., Perdew-Burke-Ernzerhof DFT, Meta-GGA DFT, experimental results).

- **Realism and complexity.** We extend model expressiveness and ability to handle real-world systems such as disordered structures, doped materials, and adsorption on complex surfaces.

In a GNN, each system is described by atomic numbers $S \in \mathbb{N}^n$ and 3D coordinates $\vec{X} \in \mathbb{R}^{n \times 3}$. A GNN learns a function $\phi : (\vec{X}, S) \rightarrow Y$ that maps the structure to a target property. Through message passing, node embeddings are iteratively updated by aggregating information from neighboring atoms, capturing both local and long-range effects. Recent developments, such as the E(3)-equivariant models MACE-MP-0 [Batatia et al., 2023] and EquiformerV2 [Liao et al., 2023], have pushed model performance to near-DFT levels, improving generalization and allowing modeling of larger systems over longer timescales.

Overall, those models enable rapid estimation of key physical properties across millions of candidates, narrowing the design space and surfacing the most promising materials for experimental validation.

2.2 Generative Models

While predictive models tell us how well a material might perform, the more fundamental challenge is knowing what to test in the first place. This is the challenge of **inverse design** and **hypothesis generation**, central to materials discovery. The space of possible compounds is incredibly large, spanning compositions, symmetries, and atomic configurations. Exploring it exhaustively is computationally infeasible, and relying solely on expert intuition limits innovation.

To tackle this, **we develop generative models** that automate the creation of chemically valid, diverse, and high-performing candidates. These models learn from data to propose materials that satisfy physical constraints and application-specific goals, enabling a scalable, systematic, and domain-aware search across the chemical space. Our core generative framework is built on **Generative Flow Networks (GFlowNets)** [Bengio et al., 2021], a class of models developed by Turing Award laureate (and Entalpic advisor) Yoshua Bengio and collaborators at Mila, Québec’s AI Institute. GFlowNets are designed to sequentially sample complex structured objects, like molecules or crystals, optimizing a reward function that encodes their utility for a given task (e.g. a geometric GNN predicting material stability).

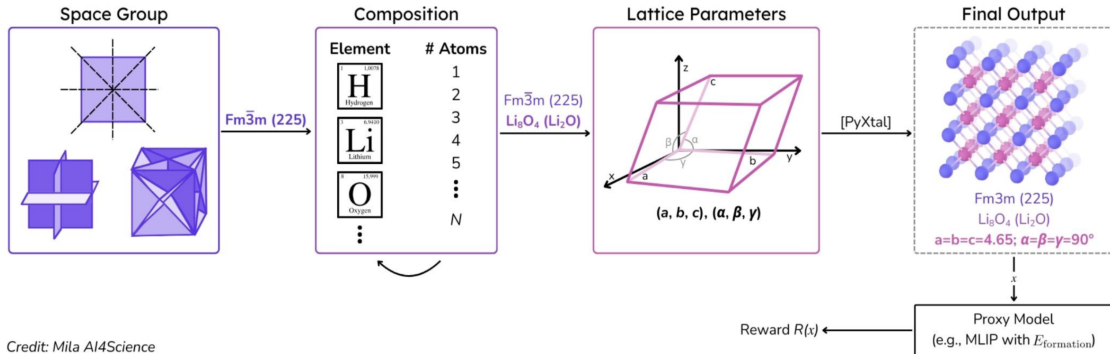


Figure 3 Simplified schematic of crystal generation process in Crystal-GFN [AI4Science et al., 2023].

Unlike diffusion models [Zeni et al., 2023], which produce candidates in a single pass, GFlowNets are inspired by **Reinforcement Learning and generate materials sequentially**. They learn a policy neural network that defines transition probabilities between partial states, thereby parametrizing an agent that chooses what sequence of actions to take in order to generate a meaningful material or molecule. This policy is trained so that the probability of generating a complete structure is proportional to the reward function. This proportionality is what enables GFlowNets to generate new and diverse structures, all-the-while jointly targeting multiple properties across different modes of the search space.

At Entalpic, we apply this framework to crystalline materials, building on top of the open-source **Crystal-GFN** model [AI4Science et al., 2023] —and extend it to catalysts and liquid chemistry. Material candidates are assembled in stages: selecting a space group, picking atomic species and assigning them to Wyckoff positions, generating lattice parameters and creating a surface from this bulk. This interpretable and controllable construction process enables us to inject physical constraints (e.g., symmetries, charge neutrality) at each decision point, guiding the model toward viable and high-performing regions of the material space.

In short, this approach provides key advantages:

- We generate diverse and novel candidates due to this step-by-step approach that samples proportionally to the reward distribution.
- We generate ‘valid’ materials, as we hard code physico-chemical rules (realistic compositions, pairwise distances, symmetries, atomic densities) to constrain the actions allowed by the policy function.
- We can control the type of structure that we explore very easily, even at inference time, allowing us to limit our search to only sensible families.
- We can optimize multiple objectives at once: target properties, synthesizability, material cost, etc.

2.3 LLMs for Chemical Prior Knowledge

Our generative models enable exploration of atomic-level design possibilities at scale, but how can we start smaller, and not just broader? Knowing what has already been tested, patented, or synthesized —and where the gaps and opportunities lie —would fundamentally shift how we navigate chemical space.

Fortunately, this knowledge already exists, scattered across decades of scientific papers and patents. With the right tools, we can **replicate and scale the intuition of leading chemists** in each field, guiding our generative models to explore the most meaningful corners of the chemical space.

At Entalpic, we therefore leverage **Large Language Models (LLMs)** with **Retrieval-Augmented Generation (RAG)** to parse this corpus and extract human-validated priors: relevant material classes, synthesis strategies and known constraints. These insights are labeled, structured, and used upstream to define application-specific search spaces for generation. In doing so, LLMs turn scientific knowledge into an actionable insights and shift the starting point —from random exploration to informed discovery —focused on the frontier between what is known and what is possible.

Together, this hybrid stack combining GNNs, GFlowNets and LLMs forms a scalable, iterative discovery engine that mirrors the scientific method of hypothesis testing, but at scale.

2.4 Active Learning : Closing the Feedback Loop

With a smart AI-driven discovery framework in place, the next challenge is clear: what should we test next? And why? Not all generated candidates are equally valuable. Some are strong contenders for real-world validation; others, if chosen wisely, offer the greatest potential to improve the models themselves and steer the discovery process forward.

This is where **active learning (AL)** becomes essential. AL helps us strike the right balance between candidates that are most likely to succeed in the lab, and those that will most effectively refine our predictive and generative models. It introduces a systematic, data-driven approach to selecting the most informative structures for high-fidelity labeling —whether through quantum simulations (e.g. Density Functional Theory [Kohn et al., 1996]) or experimental validation. Rather than treating all candidates equally, AL prioritizes those that will *shift our understanding* and sharpen our discovery engine. It is what turns our process into a continuous improvement cycle —closing the loop between simulation, prediction, and (experimental) validation.

Our active learning loop operates in three stages:

1. **Candidate generation:** The generative model proposes new structures, guided by predicted properties and uncertainty estimates.
2. **Selective evaluation:** Acquisition functions identify which candidates are most informative—for instance, those in high-uncertainty or under-explored regions of the chemical space—and forward them to high-throughput DFT or experiments pipelines.
3. **Model refinement:** Newly labeled data are integrated back into the training set, improving both prediction accuracy and generalization of our ML models.

To prioritize candidates, we use both **model-based uncertainty** (e.g., ensembles, Monte-Carlo dropout) and **data-based uncertainty** (e.g., out-of-distribution detection in an embedding latent space). These signals are incorporated in two ways: (1) directly as reward functions during GFlowNet training to guide generative exploration, and (2) as filters in post-processing to select high-value validation targets.

Compared to brute-force screening, this adaptive approach reduces cost, focuses learning where it matters most, and accelerates convergence toward high-quality materials. This feedback loop forms a complete scientific discovery iteration: *propose, test, learn, refine*. And it sets the stage for the last critical step: grounding our models in physical reality.

2.5 A Word on Experimental Validation

While AI models can screen chemical space and prioritize candidates, **only the lab can tell us what truly works**. Experimental validation thus remains the gold standard for confirming whether a promising structure is actually synthesizable, stable, and high-performing under real-world conditions.

At Entalpic, we integrate experimental workflows to close the loop between computational predictions and practical materials discovery. This stage ensures that our AI-generated hypotheses aren’t just theoretically compelling—but also experimentally sound.

This workflow consists of three tightly coupled stages:

- **Synthesis:** We collaborate with high-throughput experiments (HTE) labs, such as [Realcat](#), to enable rapid and reproducible creation of candidate materials. These partnerships allow us to efficiently validate AI-generated hypothesis through advanced synthesis capabilities.
- **Characterization:** We employ advanced analytical techniques such as X-ray diffraction (XRD), X-ray fluorescence (XRF) and electron microscopy (SEM/TEM) to understand what we synthesised exactly, probing structural and compositional fidelity.
- **Performance testing:** Materials are evaluated in application-relevant conditions using custom testing protocols and reactors, measuring properties such as catalytic activity, selectivity, and stability.

This stage does more than confirm predictions: it generates **ground truth data** that feeds back into our AI models via the active learning loop, sharpening their accuracy and surfacing failure modes that simulations alone can’t always capture—like defects, kinetics, or complex synthesis barriers.

By tightly integrating predictive models, generative design, and experimental validation through active learning, we build a virtuous cycle—one where each experiment strengthens the next hypothesis, and every iteration brings us closer to real-world breakthroughs.

3 Industrial Applications

At Entalpic, we believe that impactful innovation starts with real-world needs. That’s why we root our developments in industrial applications that reflect the most pressing challenges across sectors, from energy storage to catalysis and specialty chemistries.

Entalpic’s platform is designed to address a broad range of chemistries across three primary domains: **bulk materials, surfaces, and molecular systems**. These domains correspond directly to industrial use cases, from electrochemical devices and catalytic processes to advanced coatings or dopings.

In bulk materials, our models are applied to the discovery and optimization of solid-state systems such as battery electrodes, including doped variants and alternative chemistries like **NaFePO₄ (NFPP)**, as well as structural **aluminum-based alloys (Al-X)** for lightweight, high-strength applications.

In the **surface domain**, we explore catalytically active interfaces for key reactions such as **ammonia (NH₃) cracking, Hydrogen Evolution Reaction (HER), CO₂ reduction reactions (CO₂RR)**, and surfaces synthesized via techniques like **atomic layer deposition (ALD)**. These surface-specific use cases require precise control over atomic arrangements, which our generative models and property predictors are well-suited to handle.

On the **molecular front**, we are targeting functional small molecules such as **electrolytes or solvents**—areas where chemical structure directly impacts performance, stability, and environmental compatibility.

The applications demonstrate the breadth of use cases Entalpic can address, from electrochemistry, catalysis and specialty molecule development.

Use Case Example

Entalpic’s platform is particularly well-suited for accelerating the discovery of optimized electrode materials. One compelling application involves the design of doped NFPP ($\text{Na}_4\text{Fe}_3(\text{PO}_4)_2(\text{P}_2\text{O}_7)$) cathodes, a promising class of materials for sodium-ion batteries thanks to their structural robustness, low cost, and favorable environmental profile. However, the electrochemical performance can be further improved through targeted **doping strategies**.

Using our generative AI pipeline, Entalpic conducts a systematic exploration of substitutional dopants aimed at enhancing three critical metrics: **operating voltage, reversible capacity, and structural/diffusion stability**. Our models generate and rank dozens of chemically valid doped structures, informed by formation energy calculations and application-specific constraints such as ionic transport and electronic conductivity.

The discovery process typically begins with high-throughput screening of **single dopants** and can be extended to **co-doping strategies** in later stages. Candidates are prioritized through an active learning loop that refines both predictive and generative models based on simulation feedback. The top results are advanced toward experimental validation in collaboration with synthesis and testing partners.

This use case demonstrates how Entalpic’s platform can efficiently narrow a vast chemical space to a focused set of high-potential materials, transforming a months-long trial-and-error process into a structured, data-driven workflow that shortens discovery cycles and enhances industrial readiness.

4 Conclusion

At Entalpic, we believe the future of materials discovery lies in combining the best of both worlds: the speed and scale of AI with the insight and rigor of physical science. Our full-stack platform, spanning predictive modeling, generative design, LLMs, quantum simulations, experimental validation and active learning, offers a new paradigm for innovation.

By training models on massive datasets and guiding them with expert knowledge and real-world feedback, we accelerate discovery while maintaining control.

If you are in industrial R&D, this is the time to engage. Not just to adopt AI, but to shape it. To help build models that reflect real-world challenges and accelerate the development of sustainable, high-performance materials.

References

- Mila AI4Science, Alex Hernandez-Garcia, Alexandre Duval, Alexandra Volokhova, Yoshua Bengio, Divya Sharma, Pierre Luc Carrier, Michal Koziarski, and Victor Schmidt. Crystal-gfn: sampling crystals with desirable properties and constraints. *arXiv preprint arXiv:2310.04925*, 2023.
- Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M Elena, Dávid P Kovács, Janosh Riebesell, Xavier R Advincula, Mark Asta, William J Baldwin, Noam Bernstein, et al. A foundation model for atomistic materials chemistry. *arXiv preprint arXiv:2401.00096*, 2023.
- Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *Advances in Neural Information Processing Systems*, 34:27381–27394, 2021.
- Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, Weihua Hu, et al. Open catalyst 2020 (oc20) dataset and community challenges. *Acs Catalysis*, 11(10):6059–6072, 2021.
- Chi Chen, Dan Thien Nguyen, Shannon J. Lee, Nathan A. Baker, Ajay S. Karakoti, Linda Lauw, Craig Owen, Karl T. Mueller, Brian A. Bilodeau, Vijayakumar Murugesan, and Matthias Troyer. Accelerating computational materials discovery with artificial intelligence and cloud high-performance computing: from large-scale screening to experimental validation. *arXiv preprint arXiv: 2401.04070*, 2024.
- Juan J de Pablo, Nicholas E Jackson, Michael A Webb, Long-Qing Chen, Joel E Moore, Dane Morgan, Ryan Jacobs, Tresa Pollock, Darrell G Schlom, Eric S Toberer, et al. New frontiers for the materials genome initiative. *npj Computational Materials*, 5(1):41, 2019.
- Alexandre Duval, Simon V. Mathis, Chaitanya K. Joshi, Victor Schmidt, Santiago Miret, Fragkiskos D. Malliaros, Taco Cohen, Pietro Lio, Yoshua Bengio, and Michael Bronstein. A hitchhiker’s guide to geometric gnns for 3d atomic systems. *arXiv preprint arXiv: 2312.07511*, 2023.
- Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin a. Persson. The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 2013. ISSN 2166532X. doi: 10.1063/1.4812323. URL <http://link.aip.org/link/AMPADS/v1/i1/p011002/s1&Agg=doi>.
- Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolution networks. *arXiv preprint arXiv: 1609.02907*, 2016.
- Walter Kohn, Axel D Becke, and Robert G Parr. Density functional theory of electronic structure. *The Journal of Physical Chemistry*, 100(31):12974–12980, 1996.
- Yi-Lun Liao, Brandon Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. *arXiv preprint arXiv:2306.12059*, 2023.

- Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023.
- Martin Siron, Inel Djafar, Etienne du Fayet, Amandine Rossello, Ali Ramlaoui, and Alexandre Duval. Lemat-bulk: aggregating, and de-duplicating quantum chemistry materials databases. In *AI for Accelerated Materials Design-ICLR 2025*, 2025.
- Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, et al. Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023.
- C Lawrence Zitnick, Lowik Chanussot, Abhishek Das, Siddharth Goyal, Javier Heras-Domingo, Caleb Ho, Weihua Hu, Thibaut Lavril, Aini Palizhati, Morgane Riviere, et al. An introduction to electrocatalyst design using machine learning for renewable energy storage. *Preprint arXiv:2010.09435*, 2020.