Accelerating AI Co-scientists with HPC Infrastructure

Surya Appana
University of California, Berkeley
Berkeley, California, USA

Alok Kamatar, J. Gregory Pauloski, and Kyle Chard (Advisors)
University of Chicago
Chicago, Illinois, USA

ABSTRACT

We present MOFAI, an agentic AI Scientist coupled with high-performance computing (HPC) resources for the generation and property prediction of metal-organic frameworks (MOFs). MOFAI exhibits autonomous agents to enable tool-calling of linker generation, MOF assembly, molecular dynamics simulation, and deep learning-based property prediction in asynchronous and distributed fashion. MOFAI demonstrates success in leveraging mult-node computation to progressively discover stable MOFs with high CO_2 adsorption by learning from past successful MOFs. This seamless fusion of agentic reasoning with HPC demonstrates a new paradigm for automated scientific discovery, where AI scientists dynamically coordinate computationally intensive, domain-specific tools.

1 INTRODUCTION

Agentic systems based on large language models (LLMs) such as Google's AI Co-Scientist [2] have shown promise in automated scientific discovery. By equipping agentic systems with the ability to run in distributed environments and scale experiments on HPC infrastructure, we can further accelerate computational science. We consider an important use case in materials science: the discovery and property-prediction of Metal-Organic Frameworks (MOFs). MOFs are highly tunable porous materials ideal for applications such as carbon capture, but their discovery is hampered by a combinatorially vast chemical search space and the need to run computationally expensive simulations. Recent LLM-based approaches like the tool-augmented ChatMOF [3] and knowledge-driven dZiner [1] have shown promise in automating inverse design with the use of deep learning based surrogate models for property prediction.

We introduce MOFAI, an agentic AI Scientist that leverages the Academy [5] agentic middleware to create a seamless, dynamic coupling between autonomous LLM agents and the predictive power of HPC simulations. MOFAI's agents autonomously call linker generation, MOF assembly, and property prediction tools in an asynchronous and distributed fashion, synthesizing flexible agentic reasoning with high-fidelity computation to progressively discover stable MOFs with high CO2 adsorption.

SC25, November 18-20, 2025, St. Louis, MO

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

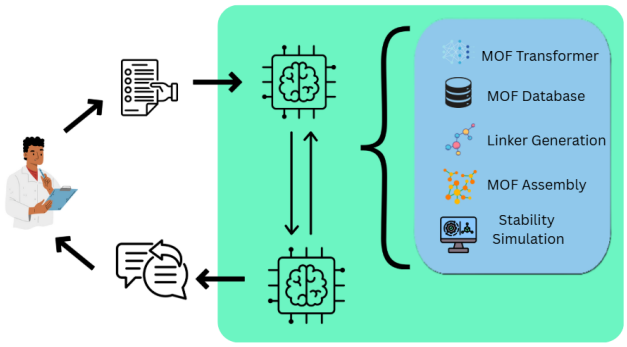


Figure 1: HPC-coupled AI Co-scientist workflow: a human scientist provides a request to the tool-calling agent, which makes use of a set of HPC-coupled tools to provide a response. This response is evaluated by a reflection agent, which either returns the response to the scientist or prompts the tool-caller to retry.

2 METHODS

We developed MOFAI for the discovery and property prediction of MOFs using Academy. Academy provides a framework to launch and coordinate stateful agents on distributed research infrastructure. This enables the MOFAI to leverage heterogeneous resources from HPC clusters to run compute intensive simulation tasks to cloud GPU nodes for model inference. MOFAI is composed of three core agent classes:

(1) **LLM-based Tool-Calling Agent** – interprets input objectives, selects tools to execute in a particular order, and formulates a response.
(2) **LLM-based Reflection Agent** – evaluates the response of the tool-calling agent and either re-prompts the agent with feedback or returns the response if it is appropriate.
(3) **Toolbox** – a set domain-specific tool agents for linker generation, MOF assembly, validation, property prediction, and MOF database querying.

The **Toolbox** integrates the following agents:

- **Prediction** – uses MOFTransformer [4], a pretrained deep learning model for predicting MOF properties from structural descriptors.
- **Database Query** – enables retrieval of example MOFs from databases, including hMOF, CoREMOF, QMOF.
- **Generation** – generates SMILES strings of candidate linkers.
- **Assembly** – combines generated linkers with inorganic metal nodes into candidate MOFs.

Table 1: An overview of the workflow agents, codes, and computational resources.

Agent	Description	Code	Resource
Generation	Generate linker SMILES	LLM APIs	Cloud
Assembly	Connect linkers & metal nodes	Custom	1 CPU, multiple nodes
Validation	Check geometry	cif2lammps	1 CPU, multiple nodes
	Test stability	LAMMPS	1 CPU, multiple nodes
Prediction	Estimate CO_2 adsorption	MOF Transformer	1 NVIDIA RTX 6000 GPU

- **Validation** – performs a LAMMPS molecular dynamics simulation with the UFF4MOF force field to test structural stability under ambient conditions.

Agents operate concurrently, exchanging results via Academy's asynchronous messaging. The MOFTransformer agent is run on a Chameleon compute node with a NVIDIA RTX 6000 GPU; the generation, assembly, and simulation agents are run on the Anvil supercomputer; and the other agents are run locally.

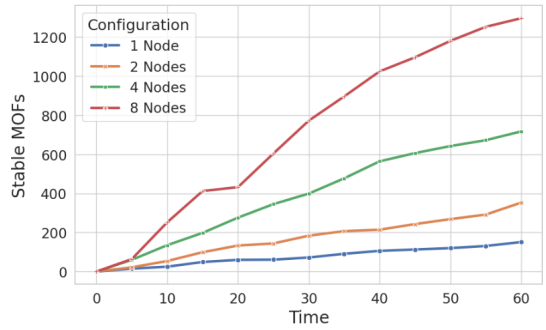
3 EVALUATION

We evaluated MOFAI's ability to automate and scale a workflow for MOF discovery optimized specifically for CO2 adsorption by enabling the assembly and validation agents to use multiple Anvil nodes. Stable MOFs are defined as MOFs with under 10% structural strain. Figure 2a demonstrates linear scaling in the number of stable MOFs generated across nodes when incorporating o3-mini for the linker generation agent and GPT-4o for the tool-calling and reflection agents. Allowing higher throughput in the screening tool calls using HPC allows for higher quality MOFs to be generated by MOFAI. The generation process was organized into rounds, and after each round we provided feedback to the generation agent on the adsorption, stability, and linkers of MOFs generated in previous rounds. Figure 2b demonstrates that CO_2 adsorption among the most recently generated stable MOFs was highest when MOFAI had access to the most nodes.

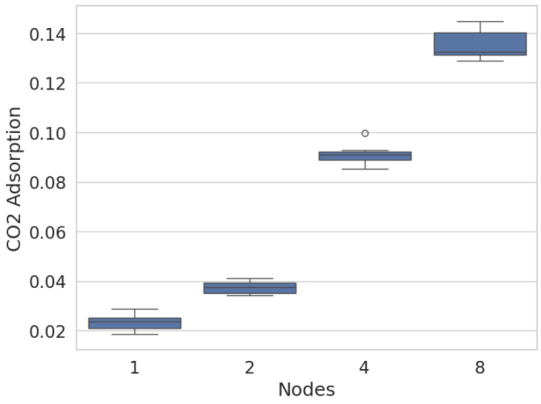
We further evaluated MOFAI's ability to generate stable MOFs across different LLMs (GPT-4o, o3-mini, Claude Sonnet 4, and Deepseek V3) used for linker generation, as shown in Figure 3, fixing GPT-4o for the tool call and reflection agents. When asked to generate 300 potential linkers, o3-mini results in all linkers being successfully assembled with default nodes into MOFs as shown in Figure 3a. Of the successfully assembled MOFs, Deepseek V3 has the highest success rate (~ 38%) in providing stable MOFs.

4 CONCLUSION

We have presented MOFAI, an agentic HPC-coupled AI Scientist for discovery and property-prediction of metal-organic frameworks (MOFs). Through remotely deployed agents for tool-calling, reflection, linker generation, MOF assembly, molecular dynamics simulation, and property prediction, MOFAI demonstrates promising results for the ability of agentic LLM systems to autonomously run

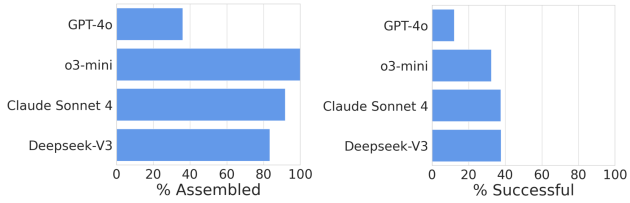


(a) Number of MOFs screened.



(b) CO2 Adsorption (mol/Kg · Pa) of returned MOFs

Figure 2: Effect of scaling AI-Scientist to multiple nodes.



(a) MOF assembly success rate **(b) Percentage of assembled MOFs that are stable.**

Figure 3: MOF assembly and LAMMPS success across linker generation LLMs when generating 300 linkers.

HPC-coupled tools in the process of MOF discovery. Future research efforts will hopefully expand on this ability to more seamlessly pair AI Scientists with complex tools in scientific domains.

REFERENCES

[1] Mehrad Ansari, Jeffrey Watchorn, Carla E. Brown, and Joseph S. Brown. 2024. dZiner: Rational Inverse Design of Materials with AI Agents. arXiv:2410.03963 [physics.chem-ph] https://arxiv.org/abs/2410.03963

[2] Juraj Gottweis, Wei-Hung Weng, Alexander Daryin, Tao Tu, Anil Palepu, Petar Sirkovic, Artiom Myaskovsky, Felix Weissenberger, Keran Rong, Ryutaro Tanno, Khaled Saab, Dan Popovici, Jacob Blum, Fan Zhang, Katherine Chou, Avinatan Hassidim, Burak Gokturk, Amin Vahdat, Pushmeet Kohli, Yossi Matias, Andrew Carroll, Kavita Kulkarni, Nenad Tomasev, Yuan Guan, Vikram Dhillon, Eeshit Dhaval Vaishnav, Byron Lee, Tiago R D Costa, José R Penadés, Gary Peltz, Yunhan Xu,

Annalisa Pawlosky, Alan Karthikesalingam, and Vivek Natarajan. 2025. Towards an AI co-scientist. arXiv:2502.18864 [cs.AI] https://arxiv.org/abs/2502.18864

[3] Yeonghun Kang and Jihan Kim. 2023. ChatMOF: An Autonomous AI System for Predicting and Generating Metal-Organic Frameworks. arXiv:2308.01423 [cs.CL] https://arxiv.org/abs/2308.01423

[4] Yeonghun Kang, Hyun Park, Berend Smit, and Jihan Kim. 2022. MOFTransformer: A Multi-modal Pre-training Transformer for Universal Transfer Learning in Metal-Organic Frameworks. *ChemRxiv* (2022). https://doi.org/10.26434/chemrxiv-2022-hcjzc

[5] J. Gregory Pauloski, Yadu Babuji, Ryan Chard, Mansi Sakarvadia, Kyle Chard, and Ian Foster. 2025. Empowering Scientific Workflows with Federated Agents. arXiv:2505.05428 [cs.MA] https://arxiv.org/abs/2505.05428