

Research Highlight

Open Access



Artificial intelligence-driven autonomous laboratory for accelerating chemical discovery

Junwu Chen^{1,*} , Qiucheng Xu^{2,*}

¹Laboratory of Artificial Chemical Intelligence (LIAC), Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland.

²Laboratory of Inorganic Synthesis and Catalysis, Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland.

* **Correspondence to:** Junwu Chen, Laboratory of Artificial Chemical Intelligence (LIAC), Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland. E-mail: junwu.chen@epfl.ch; Dr. Qiucheng Xu, Laboratory of Inorganic Synthesis and Catalysis, Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland. E-mail: qiucheng.xu@epfl.ch

How to cite this article: Chen, J.; Xu, Q. Artificial intelligence-driven autonomous laboratory for accelerating chemical discovery. *Chem. Synth.* **2025**, *5*, 76. <https://dx.doi.org/10.20517/cs.2025.66>

Received: 22 May 2025 **First Decision:** 15 Jul 2025 **Revised:** 24 Jul 2025 **Accepted:** 31 Jul 2025 **Published:** 18 Sep 2025

Academic Editor: Guangshan Zhu **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Keywords: Artificial intelligence, machine learning, autonomous laboratory, chemical synthesis

MAIN TEXT

Autonomous laboratories, also known as self-driving labs, have emerged as a powerful strategy to accelerate chemical discovery^[1-3]. By highly integrating different key parts including artificial intelligence (AI), robotic experimentation systems and automation technologies into a continuous closed-loop cycle, autonomous laboratories can efficiently conduct scientific experiments with minimal human intervention^[4-6]. In particular, AI plays a central role in key stages such as experimental planning, synthesis recipe design and optimization, as well as data analysis and interpretation in characterization techniques. In an ideal case, given a target molecule or material, the AI model trained on literature data and prior knowledge generates initial synthesis schemes, including precursors, intermediates for each step, and reaction conditions. Robotic systems then automatically carry out every step of the synthesis recipe, from reagent dispensing and reaction control to sample collection and product analysis. The characterization data of the product is analyzed by software algorithms or machine learning (ML) models for substance identification and yield estimation, based on which improved synthetic routes are proposed with the assistance of AI techniques



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



such as active learning and Bayesian optimization. This closed-loop approach minimizes downtime between manual operations, eliminates subjective decision points, and enables rapid exploration of novel materials and optimization strategies. By tightly integrating these stages (i.e., protocol design, hands-off execution, and data-driven learning), autonomous labs aim to turn processes that once took months of trial and error into routine high-throughput workflows.

Early in 2023, Szymanski *et al.* built A-Lab, a fully autonomous solid-state synthesis platform powered by AI tools and robotics [Figure 1A]^[3]. In this demonstration, they integrated four key components: (1) selection of novel and theoretically stable materials using large-scale *ab initio* phase-stability databases from the Materials Project and Google DeepMind; (2) synthesis recipe generation via natural-language models trained on the literature data; (3) phase identification from X-ray diffraction (XRD) patterns via ML models; and (4) active-learning driven optimization of synthesis routes. Over 17 days of continuous operation, A-Lab synthesized 41 of 58 DFT-predicted, air-stable inorganic materials, achieving a 71% success rate with minimal human intervention. Central to its performance were ML models for precursor and synthesis temperature selection, convolutional neural networks for XRD phase analysis, and the ARROWS3 algorithm for iterative route improvement. This work demonstrates that autonomous materials discovery at scale is feasible and points the way toward self-driving laboratories that can accelerate materials innovation.

Several months ago, Dai *et al.* demonstrated a modular autonomous platform for exploratory synthetic chemistry by integrating free-roaming mobile robots with standard laboratory instruments and a heuristic decision maker [Figure 1B]^[1]. In their setup, mobile robots sample, transport, and operate a Chemspeed ISynth synthesizer, an ultraperformance liquid chromatography (UPLC)–mass spectrometry (MS) system, and a benchtop nuclear magnetic resonance (NMR) spectrometer, all coordinated by heuristic decision maker that processes orthogonal analytical data to mimic expert judgments. In detail, the heuristic reaction planner assigns pass or fail to both MS and NMR results, using dynamic time warping to detect reaction-induced spectral changes and a precomputed *m/z* lookup table, before determining the next experimental steps. By applying human-like criteria rather than optimizing a single metric, the system autonomously performs screening, replication, scale-up, and functional assays over multi-day campaigns, exploring complex chemical spaces such as structural diversification chemistry, supramolecular assembly, and photochemical catalysis. This approach accelerates reaction discovery through instantaneous decision making and shared instrumentation, offering a scalable blueprint for broadly accessible self-driving chemistry laboratories.

Recent advances in large language models (LLMs) have also rapidly expanded both the capabilities and applications of autonomous laboratories. Several pioneering studies have demonstrated the potential of LLMs-based agents to serve as the “brain” of autonomous chemical research^[5,7]. Boiko *et al.* proposed Coscientist, an LLM-driven system capable of autonomously designing, planning, and controlling robotic operations for chemical experiments^[7]. The LLM agent in Coscientist is equipped with tool-using capabilities that enable it to perform tasks such as web searching, document retrieval, code generation and execution, as well as interaction with robotic experimentation systems. Coscientist demonstrated its promise to accelerate research in six chemistry tasks, including the successful optimization of palladium-catalyzed cross-coupling reactions. Simultaneously, Bran *et al.* developed ChemCrow, an LLM agent that integrates 18 expert-designed tools to enhance chemical research capabilities^[5]. Augmented by external tools, ChemCrow can design and plan experiments, and autonomously perform complex chemical tasks that are beyond the capabilities of conventional LLMs. For example, in the task of synthesizing an insect repellent or designing a thiourea organocatalyst of Diels–Alder reaction, ChemCrow sequentially utilizes tools to find suitable molecules, plan synthetic routes, and execute the synthesis on the cloud-based robotic

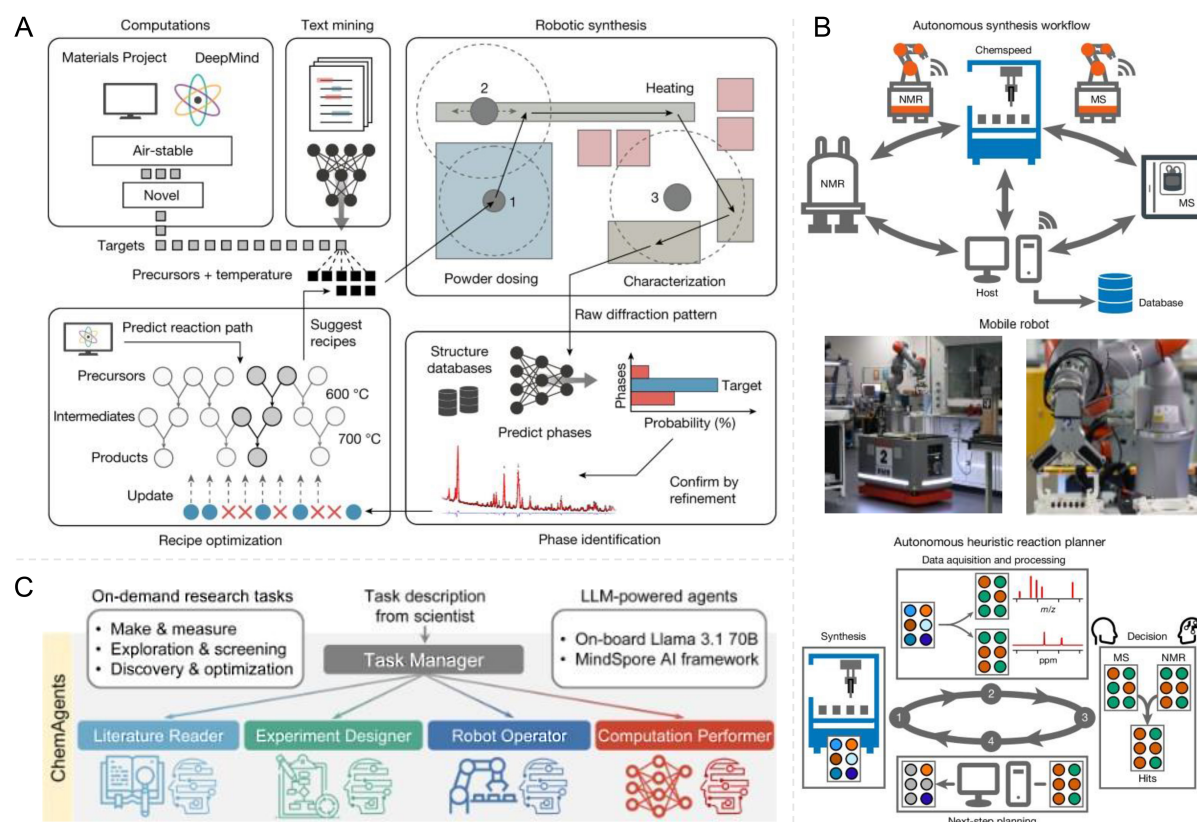


Figure 1. (A) Workflow of A-Lab for autonomous materials discovery, integrating ab initio target selection, ML-driven synthesis recipe generation, robotic solid-state synthesis, ML-driven phase identification, and active-learning optimization^[3]. Copyright 2023 Springer Nature; (B) Modular robotic workflow with mobile robots (digital images) transporting samples between a Chemspeed ISynth synthesizer, UPLC-MS and benchtop NMR, guided by a heuristic reaction planner^[1]. Copyright 2024 Springer Nature; (C) ChemAgents: a LLM-based hierarchical multi-agent system featuring a central Task Manager that coordinates four role-specific agents (Literature Reader, Experiment Designer, Computation Performer, Robot Operator) for on-demand autonomous chemical research^[4]. Copyright 2025 American Chemical Society. ML: Machine learning; UPLC-MS: ultraperformance liquid chromatography-mass spectrometry; NMR: nuclear magnetic resonance; LLM: large language model.

automation experiment platform from IBM Research. Recently, Song *et al.* introduced ChemAgents [Figure 1C]^[4], an on-demand autonomous chemical research system built on a hierarchical multi-agent architecture powered by open source LLMs. By integrating tool-calling capabilities into this LLM-driven framework, ChemAgents can interpret natural-language research prompts, generate detailed protocols, execute multistep experiments, and iteratively optimize outcomes with minimal human oversight. Its capabilities were demonstrated across seven “on-demand” experimental tasks: in “Make and Measure” tasks, it automated synthesis and Fourier transform infrared (FTIR) analysis of azobenzene derivatives, powder X-ray diffraction (PXRD) characterization of metal oxides, and fluorescence spectroscopy of perovskite quantum-dot films; in “Exploration and Screening” tasks, it conducted full-factorial optimization of graphitic carbon nitride for hydrogen evolution and screened bismuth oxyhalides for tetracycline photodegradation; in “Discovery and Optimization” tasks, it employed Bayesian optimization to identify a high-entropy metal-organic catalyst for the oxygen evolution reaction with low overpotential and exceptional stability; and in “Portability and Adaptability” tasks, it was redeployed in a different laboratory to autonomously perform a photocatalytic debromination. This hierarchical multi-agent approach paves the way for adaptable, on-demand self-driving laboratories that can plan, execute, and refine experiments across diverse chemical domains, democratizing high-throughput discovery and dramatically accelerating research.

Although autonomous laboratories hold great promise and have developed rapidly in recent years, they still face several key constraints that limit their widespread deployment and effectiveness. First, the performance of AI models depends heavily on high-quality and diverse data. However, experimental data often suffer from data scarcity, noise, and inconsistent sources, which could hinder AI models from accurately performing tasks such as materials characterization, data analysis, and product identification. Moreover, most autonomous systems and AI models are highly specialized for specific reaction types, materials systems, or experimental setups. AI models often struggle to generalize across different domains or conditions, which limits their transferability to new scientific problems. Regarding LLM-based decision-making, LLMs could generate plausible but incorrect chemical information, including impossible reaction conditions or incorrect references and data. LLMs often provide confident-sounding answers without indicating uncertainty levels, potentially leading to expensive failed experiments or safety hazards when operating outside their training domains. Hardware constraints also hinder the generalization of autonomous laboratories. Different chemical tasks require different instruments, e.g., solid-phase synthesis requires furnaces, powder handling, and XRD, while organic synthesis requires liquid handling and NMR. Current platforms lack modular hardware architectures that can seamlessly accommodate diverse experimental requirements. In addition, autonomous laboratories may misjudge or crash when faced with unexpected experimental failures, outliers, or new phenomena. Robust error detection, fault recovery, and adaptive planning remain underdeveloped.

In summary, these platforms demonstrate that autonomous labs can dramatically accelerate chemical synthesis and materials innovation by seamlessly integrating design, execution, and optimization into a self-driven cycle. Looking ahead, boosting the intelligence, capacity, and reliability of autonomous labs will require integrating more advanced AI models, leveraging reinforcement learning for adaptive control, and adopting cloud-based platforms for collaborative experimentation. To enhance the generalization of AI models, it is necessary to train foundation models or domain-adaptive models across different materials and reactions, and use methods such as transfer learning and meta-learning to adapt the model to limited new data. To overcome hardware constraints, it is necessary to develop standardized interfaces that allow rapid reconfiguration of different instruments, and extend the mobile robot capabilities to include specialized analytical modules that can be deployed on demand. To address the problems of scarcity and inconsistent sources of model training data, it is necessary to develop standardized experimental data formats, utilize high-quality simulation data and uncertainty analysis. Moreover, embedding targeted human oversight during development will streamline error handling and strengthen quality control.

DECLARATIONS

Authors' contributions

Wrote the manuscript: Chen, J.; Xu, Q.

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

Both authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2025.

REFERENCES

1. Dai, T.; Vijayakrishnan, S.; Szczypiński, F. T.; et al. Autonomous mobile robots for exploratory synthetic chemistry. *Nature* **2024**, *635*, 890-7. DOI PubMed PMC
2. Gao, W.; Raghavan, P.; Coley, C. W. Autonomous platforms for data-driven organic synthesis. *Nat. Commun.* **2022**, *13*, 1075. DOI PubMed PMC
3. Szymanski, N. J.; Rendy, B.; Fei, Y.; et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **2023**, *624*, 86-91. DOI PubMed PMC
4. Song, T.; Luo, M.; Zhang, X.; et al. A multiagent-driven robotic AI chemist enabling autonomous chemical research on demand. *J. Am. Chem. Soc.* **2025**, *147*, 12534-45. DOI PubMed
5. Bran, A. M.; Cox, S.; Schilter, O.; Baldassari, C.; White, A. D.; Schwaller, P. Augmenting large language models with chemistry tools. *Nat. Mach. Intell.* **2024**, *6*, 525-35. DOI PubMed PMC
6. Zhao, Y.; Zhao, Y.; Wang, J.; Wang, Z. Artificial intelligence meets laboratory automation in discovery and synthesis of metal-organic frameworks: a review. *Ind. Eng. Chem. Res.* **2025**, *64*, 4637-68. DOI
7. Boiko, D. A.; MacKnight, R.; Kline, B.; Gomes, G. Autonomous chemical research with large language models. *Nature* **2023**, *624*, 570-8. DOI PubMed PMC

**Junwu Chen**

Junwu Chen obtained his B.Sc. degree in Chemical Engineering from Wuhan University of Technology in 2019 under the supervision of Prof. Weiquan Cai, and his M.Sc. degree in Chemical Engineering from the University of Chinese Academy of Sciences in 2022 under the supervision of Prof. Suojiang Zhang and Prof. Kun Dong. He is currently a Ph.D. student at École Polytechnique Fédérale de Lausanne (EPFL), advised by Prof. Philippe Schwaller, working on artificial intelligence for chemistry and materials.

**Qiucheng Xu**

Dr. Qiucheng Xu is a Postdoctoral Researcher at École Polytechnique Fédérale de Lausanne (EPFL). He received his Ph.D. in 2021 from East China University of Science and Technology under the supervision of Prof. Chunhong Li and Prof. Hao Jiang. From 2021 to 2024, he conducted postdoctoral research at the Technical University of Denmark, working with Prof. Brian Seger and Prof. Ib Chorkendorff. Since April 2024, he has been a postdoctoral researcher at EPFL in Prof. Xile Hu's group. His recent research focuses on the controlled synthesis of efficient electrocatalytic materials and the failure analysis of membrane-electrode assembly (MEA)-based electrochemical devices. He is also actively exploring the integration of artificial intelligence for electrochemical data analysis and performance prediction, with the long-term goal of advancing autonomous laboratory systems.