https://doi.org/10.1038/s41578-023-00588-4

Autonomous experiments using active learning and AI

Zhichu Ren, Zekun Ren, Zhen Zhang, Tonio Buonassisi & Ju Li

Active learning and automation will not easily liberate humans from laboratory workflows. Before they can really impact materials research, artificial intelligence systems will need to be carefully set up to ensure their robust operation and their ability to deal with both epistemic and stochastic errors. As autonomous experiments become more widely available, it is essential to think about how to embed reproducibility, reconfigurability and interoperability in the design of autonomous labs.

Materials discovery is a labour-intensive process. Edison famously tested thousands of filaments to develop the incandescent light bulb. Nowadays, affordable automation is enabling the emergence of a new research modality that incorporates robotics and active learning algorithms[1,2]. Constructing fully automated experimental platforms[3,4] is challenging when budget and space are limited, but it is entirely ok to begin with a semi-automated workflow, such as with manual transfer of sample arrays between instruments. Basic active learning approaches using Gaussian process regression and Bayesian optimization and their variants can satisfactorily manage many types of process optimization, provided that the input–output pairs are reproducible[5,6].

Just as raising children takes decades and different kinds of lessons, one should not anticipate active-learning-driven experiments with a limited knowledge base to be very productive at the outset. The process is fragile at the beginning. It takes a lot of handholding and communication to teach a toddler to walk, and one should expect the same with active learning and artificial intelligence (AI)-driven experiments – much guidance is needed, even with a seemingly robust automated pipeline.

Addressing epistemic errors

The ability to obtain long-term reproducible datasets is the hallmark of a mature robotic platform qualified for carrying out active learning. When an experiment is repeated twice and produces varying outcomes, the disparity arises from two origins: aleatoric errors and epistemic errors. Aleatoric errors stem from stochasticity and are easier to handle, because they can be relieved by automation and inferred by the Gaussian process noise kernel. On the contrary, epistemic errors could wreak havoc on autonomous experiments driven by naive active learning algorithms. Essentially, what happens is that we take the constancy of some variables for granted, whereas in reality they vary. For example, in our robotic platform, for a while we found large variations in performance of carbon-paper-based samples prepared from a simple-looking automated liquid drop-casting process. This issue was not resolved until we noticed that the carbon substrate could be anisotropic, which means the way we cut it is an important variable (more details available in this case study).

One may wonder why reproducibility is particularly critical for active learning. Don't experimentalists doing manual experiments also face this issue? The answer is yes, but it is much relieved by the vast human experience and fluid intelligence. Imagine a student who discovers a synthesis recipe and repeats it 10 times, obtaining very exciting results 2 out of 10 times. What will the student do? The error bar is too big to publish the findings, so the student and the advisor will discuss, tweak the setup in many ways and eventually figure out the reason behind the statistical anomaly (which could be, for example, the extraneous moisture content of an intermediate reaction product).

Statistical anomalies stem from our inability to identify all the underlying variables that contribute to the result and, if they are ignored rather than investigated and understood, lead to irreproducibility. A *Nature* survey revealed that the primary reason for irreproducibility in the literature is selective reporting[7]. If one recklessly launches an active learning project without identifying the source of the error bars, the effort could waste a lot of time and money. The algorithm will mistakenly treat spurious noises as signals, consequently giving poor suggestions, as 'garbage in, garbage out'.

The flip side of the coin is that epistemic errors, if carefully debugged, can lead to wondrous scientific discoveries (for example, penicillin was discovered by failing to grow bacterial cultures with unintended fungal contamination). Humans are very good at turning around 'experimental failures', as we have exceptional causal inference capabilities ("once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth", to quote Sherlock Holmes). This is not the case with basic active-learning methods, because they take an over-simplified view of the world and do not have much prior physical knowledge.

Unlike conventional machine learning techniques, large language models such as ChatGPT can be used to generate well-posed scientific hypotheses[8]. In the future, it will be possible to use broader and deeper laboratory automation to experimentally test these machine-generated hypotheses, which may be able to explain epistemic errors. For example, a synthetic procedure can be automatically repeated inside various controlled atmospheric chambers to understand the dependence on the partial pressure of different gases. As automated experiments start to incorporate computer vision, which can already outperform human vision in certain tasks, the ability to track laboratory conditions (such as moisture, radiation background, precursor materials texture and non-uniformity) much more meticulously than humans, and a large prior knowledge base about how the world works materially[8], it is only a matter of time before AI systems with multi-modal sensors

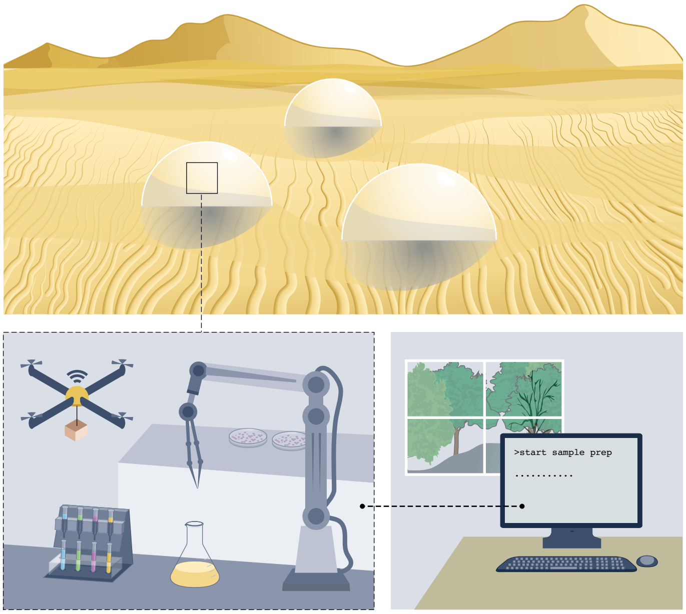

Fig. 1 | Future outlook: autonomous labs connected in an AI network. With the increasing capabilities of artificial intelligence (AI) systems and limitations in budget and physical space, the adoption of modular cloud laboratory facilities could be favourable. These could be implemented on a large scale in wastelands. These facilities allow for the flexible reconfiguration and interconnection of equipment chains, ensuring interoperability across multiple autonomous laboratories. Digit flow and mass flow are the two most critical streams: multiple AI agents interchange information via a unified network, while in the physical realm, numerous wheeled robots and drones function as the propellant for the transit of physical samples. Researchers from any corner of the globe can remotely access the system, and their commands, given in human language, may be parsed into subtasks and distributed by large language models[10].

can figure out the plausible causes of epistemic errors and debug the workflow experimentally. Large language models plus reinforcement learning with generalized sensorimotor functions and the 'new cybernetics' described below may be the next stage of the lab-automation revolution.

Towards interconnected AI-driven labs

As AI systems become more complex and powerful, budgetary and space constraints require the use of modular cloud lab facilities[9] (Fig. 1), so that the equipment chain can be 'recompiled' and relinked and the interoperability between multiple autonomous labs can be ensured. A network of AI systems, both experimental and theoretical, is needed to achieve partition of labour, an economy of scale, and checks and balances (as in round-robin tests with physical sample transfer and adversarial peer reviews[8]).

Today's commercial equipment for materials synthesis, characterization and property testing is designed with just human users in mind. In the future, an autonomous lab would require every piece of equipment to have two interfaces, a main interface for AI systems on the Internet of Things and a human access interface. Each apparatus would operate similarly to a subroutine in a software library, with physical sample input/output specifications rigorously defined. Flexible chains of modular equipment would be designed to be reconfigured (disassembled and reassembled) quickly and automatically. Note that reconfigurability does not always require apparatuses to be physically moved to form an assembly line, as wheeled robots and small flying drones could be used to transfer samples between modules.

Although autonomous materials discovery labs have been envisioned and developed since the 1950s, there have been relatively few real successes so far[1–3]. At academic institutions, the budget for building each lab is limited to a few million US dollars or less. This means one-trick or few-trick ponies, ill-prepared for the identification of epistemic errors and rapid changes of course in the workflow. Whereas human researchers suspecting something unusual can leave their comfort zone and ask colleagues to do complementary measurements just by walking between different facilities on campus, today's autonomous labs, which still tend to be too small and subcritical, do not have this flexibility yet.

To overcome this issue, autonomous labs need to work together. We need to allow AI agents to communicate with each other through universal sample transfer and data-transfer procedures. This would allow Autonomous Lab A to send physical samples to Autonomous Lab B, with the associated meta-data. Standardized capsules for transferring liquid, powder, gel, pellet and single-crystal materials need to be developed, and they need to be compatible with easy weighing, sizing, and optical and chemical characterization techniques, plus they need to prevent contamination. We will also need buildings and infrastructure designed specifically for flexible automation. Entirely new architectures can be constructed for robot and human researchers to work together.

The era of AI has approached. To fully release the potential of AI in experimental research and materials discovery, it is crucial to equip silicon-based intelligence with 'hands' (materials synthesis, equipment self-assembly/disassembly and sample transfer) and 'eyes' (materials characterization and multi-modal sensing). Building a robust AI-to-real-world feedback system is certainly not an easy job. But as AI labs will be set up and interconnected properly, and the know-how — crystallized in standardized interfaces and hardware modules — will be broadly shared worldwide, powerful AI-enabled flexible and robust experimental workflows could revolutionize materials research.

Zhichu Ren[1], Zekun Ren[2], Zhen Zhang[1], Tonio Buonassisi[2,3] ✉ & Ju Li[1,4] ✉

[1]Department of Materials Science and Engineering, MIT, Cambridge, MA, USA. [2]Xinterra, Singapore, Singapore. [3]Department of Mechanical Engineering, MIT, Cambridge, MA, USA. [4]Department of Nuclear Science and Engineering, MIT, Cambridge, MA, USA.

✉e-mail: buonassi@mit.edu; liju@mit.edu

Published online: 3 August 2023

References

1. Morgan, D. et al. Machine learning in nuclear materials research. *Curr. Opin. Solid State Mater. Sci.* **26**, 100975 (2022).
2. Coley, C. W., Eyke, N. S. & Jensen, K. F. Autonomous discovery in the chemical sciences part II: outlook. *Angew. Chem. Int. Ed.* **59**, 23414–23436 (2020).
3. Burger, B. et al. A mobile robotic chemist. *Nature* **583**, 237–241 (2020).
4. Chen, J. et al. Navigating phase diagram complexity to guide robotic inorganic materials synthesis. Preprint at https://arxiv.org/abs/2304.00743 (2023).
5. Stach, E. et al. Autonomous experimentation systems for materials development: a community perspective. *Matter* **4**, 2702–2726 (2021).
6. Siemenn, A. E., Ren, Z., Li, Q. & Buonassisi, T. Fast Bayesian optimization of needle-in-a-haystack problems using zooming memory-based initialization (ZoMBI). *npj Comp. Mater.* **9**, 79 (2023).
7. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
8. Park, Y. J. et al. Can ChatGPT be used to generate scientific hypotheses? Preprint at https://arxiv.org/abs/2304.12208 (2023).
9. Arnold, C. Cloud labs: where robots do the research. *Nature* **606**, 612–613 (2022).
10. Ren, Z. C., Zhang, Z., Tian Y. S. & Li, J. CRESt – Copilot for Real-world Experimental Scientist. Preprint at https://doi.org/10.26434/chemrxiv-2023-tnz1x (2023).

Acknowledgements

The authors thank Y. Tian for insightful discussions and R. S. Indradjaja for giving feedback on the manuscript. They acknowledge support by DTRA (award no. HDTRA1-20-2-0002) Interaction of Ionizing Radiation with Matter (IIRM) University Research Alliance (URA).

Competing interests

Zekun Ren and T.B. are co-founders of Xinterra Pte. Ltd, a startup focused on applying active learning to accelerate the development of materials for sustainability. The other authors declare no competing interests.