Please cite this article in press as: Epps et al., Universal self-driving laboratory for accelerated discovery of materials and molecules, Chem (2021), https://doi.org/10.1016/j.chempr.2021.09.004

Chem

CellPress

Synergy

Universal self-driving laboratory for accelerated discovery of materials and molecules

Robert W. Epps,[1] Amanda A. Volk,[1] Malek Y.S. Ibrahim,[1] and Milad Abolhasani[1,*]

Self-driving laboratories are quickly growing in capability, making research in the exploration of advanced functional materials and molecules on the edge of a new era of productivity. As researchers near the widespread adoption of these powerful tools, we must assess their trajectory and the impact of their future developments.

For as long as scientific methods have existed, researchers have desired faster and more efficient methods of experimentation and discovery. With the rapid rise of robotics and artificial intelligence (AI)-guided research, we are on the horizon of a renaissance in chemistry that fulfills these aspirations. Tasks that previously required extensive time, labor, and reagents can now be automatically conducted with greater precision, efficiency, and scope. As a result, researchers may focus on defining the next big scientific problem, employ more creative exploration techniques, and gain access to otherwise unreachable regions of the chemical universe.

What is a self-driving lab?

A self-driving laboratory is comprised of two components: (1) the hardware that automatically prepares the precursors, conducts the experiment, and measures the outcome; and (2) the AI *brain* (i.e., the data-driven modeling/decision-making strategy), which analyzes the data and autonomously selects the next experiment based on the pre-set objective by a human researcher. The self-driving laboratory serves as an assistant to scientists, who define its objectives as well as initial hypotheses and chemical and physical boundaries.

In existing work focusing on advanced functional materials and molecules, the hardware of self-driving laboratories takes a variety of forms. These systems include robotic platforms spanning entire labs[1] down to compact workstations for precursor preparation and sample handling.[2] Such workstations can be integrated with batch or flow reactors for automatically conducting reactions in series or parallel.[3,4]

Regarding the AI brain of self-driving laboratories, experiment selection algorithms depend primarily on the nature of the research being conducted. For example, pharmaceutical research has mostly used cheminformatic-based strategies that employ a combination of physical models and literature data to select high probability candidate molecules and reaction synthesis routes.[5] Conversely, the highly sensitive and multidimensional nature of nanomaterial syntheses has made lab-to-lab and batch-to-batch consistency difficult to achieve, hindering the broad adoption of informed AI methods to nanoscience research. Therefore, self-driving laboratories for nanoscience studies have typically relied on AI algorithms that excel without prior knowledge, such as Bayesian optimization, reinforcement learning, or evolutionary algorithms.[6]

CHEMPR 1396

Please cite this article in press as: Epps et al., Universal self-driving laboratory for accelerated discovery of materials and molecules, Chem (2021), https://doi.org/10.1016/j.chempr.2021.09.004

How to build a self-driving lab?

Self-driving experimentation platforms have achieved notable success in both academic research and industry, but application of these technologies by a non-specialized researcher comes with several challenges. First, determining the ideal hardware for experimentation is not straightforward and depends on the material or molecules of interest. Robotic systems integrated with batch reactors are the more versatile approach and correlate directly to most methods found in literature with respect to heat and mass transfer rates. They also have access to most characterization methods a human operator may use. However, these systems have slow sampling rates, consume large quantities of reagents per condition (milliliters to liters), and generally cannot operate for extended periods without user intervention. Parallelized batch reactors and more specialized combinatorial screening systems can be significantly more time and material efficient, with reagent volumes down to a nanoliter scale and sampling rates on the order of thousands per day,[7] but they have limited control of the reaction environment and access to precise online characterization methods. Flow reactors are highly efficient (microliter reagent consumption and sampling rates rivaling combinatorial screening) and can combine the benefits of a large library of *in situ* and online characterization techniques together with precise control over reaction parameters. However, the benefit of flow reactors is a double-edged sword. Their high heat and mass-transfer rates make them an ideal choice for process intensification, but the different heat and mass-transfer dynamics of flow reactors compared with batch make it more difficult for researchers to adopt literature protocols, specifically for nanomaterials, developed through batch reactions. Most critically, flow reactors struggle to accommodate solid reagents, products, and byproducts. Because of the shortcomings of each strategy outlined, completely different automated experimentation strategies are often necessary at different stages of discovery and development of advanced functional materials and molecules. Therefore, many examples of self-driving labs have been restricted to isolated reaction stages instead of covering the full experimentally accessible parameter space of a specific class of materials or molecules.

Beyond these difficulties in optimal platform selection in multi-stage systems, researchers looking to build an autonomous experimental platform must also navigate an absence of readily available equipment. Navigating hardware availability is a problem beyond complex multi-stage reactions. Designing and building a self-driving experimental system from the ground up is costly and requires a considerable time investment. This barrier is not a significant issue for researchers specialized in platform development because the design itself is the end goal of the work, but for a chemist or material scientist aiming to improve a synthesis without advancing an experimentation platform, this limitation creates a large barrier. Furthermore, without consistency in reaction environments, identical input conditions can likely result in different reaction products between two different self-driving platforms equipped with different size reactors. As the capabilities of autonomous systems grow further toward general application, it will be critical for the field to emphasize the development of systems built from accessible and standardized components that produce consistent results. However, historically, widespread adoption of standards in unregulated communities has been driven by either extreme necessity or convenience. In self-driven experimental systems, development efficiency in future studies will likely rely on the latter.

One already occurring example of convenience-driven standardization has been the use of tubing-based flow reactors.[8] In these systems, reactions are conducted in commercially available micro scale junctions and tubing channels, typically

Please cite this article in press as: Epps et al., Universal self-driving laboratory for accelerated discovery of materials and molecules, Chem (2021), https://doi.org/10.1016/j.chempr.2021.09.004

composed of chemically resistant materials (Teflon or stainless steel), and the dimensions of these channels are manufactured with high precision under standardized dimensions. Consequently, research in developing these tubular flow reactors has formed a library of readily available, high-efficiency experimentation devices with directly transferable heat- and mass-transfer characteristics between systems. Similarly, many flask-based automated systems rely on custom 3D-printed modules coupled with commercial components, both of which may be quickly reproduced and applied in new applications.[9] It is the onus of academic research to emphasize these more accessible variants of automated experimental tools and highlight the importance of accessibility.

What is next for self-driving labs?

The next critical steps in autonomous robotic experimentation toward achieving a universal self-driving lab will be (1) the introduction of greater transparency in system design, (2) the modularization and standardization of the hardware, and (3) the creation of open-access datasets for benchmarking and selecting suitable AI modeling and decision-making algorithms.

AI, computational, and machine-learning communities have long valued transparency and open resources in academic publications. Equivalent community standards would be hugely advantageous in the field of autonomous robotic experimentation. A self-driving lab capable of autonomous planning and conduction of reactions may be significant in its abilities, but it possesses little functional application to the broader scientific community if it cannot be reconstructed in a different environment. Field standards for reporting of novel autonomous platforms should, therefore, allow for the complete reproduction of the system by an uninformed, skilled scientist and include all associated control and analysis software and relevant component models. Furthermore, widespread publication of all generated experimental data with comprehensive demonstrations of sampling precision would allow for the rapid development and benchmarking of cheminformatic and materials informatics strategies for different classes of materials and molecules. The various biases of published data skew the effectiveness of literature-driven algorithms toward high-performing regions of the chemical universe, and comprehensive reporting, including *failed* reactions, would fill many of the information gaps not currently covered in the literature. For black-box algorithms, one of the current challenges in many fields is the selection of suitable algorithms for specific scenarios. Off-the-shelf decision-making algorithms often cannot be directly applied to a focused application without further tuning of the meta-decision structure and algorithm parameters. From the perspective of an experimentalist, required algorithm tuning can slow research and even defeat the purpose of applying the algorithm to begin with. Data availability would enable researchers to test new algorithms on multiple experiment-based benchmark surrogate systems, leading to expedited implementation of higher performing algorithms.

Although autonomous robotic experimentation will most likely not converge onto a single optimal design, researchers today can make significant gains in shifting the field from its current state of diverse, isolated platforms toward a single unified system of modularized, self-driving labs. No single field in autonomous experimentation possesses all the tools necessary to explore the complexities of the chemical world, but a combined approach could bring the scientific community much closer. Modularization of experimental systems enables cross-disciplinary application of otherwise inaccessible devices and tools. A publicly available library of accessible platform designs with corresponding control and experiment selection algorithms,

Please cite this article in press as: Epps et al., Universal self-driving laboratory for accelerated discovery of materials and molecules, Chem (2021), https://doi.org/10.1016/j.chempr.2021.09.004

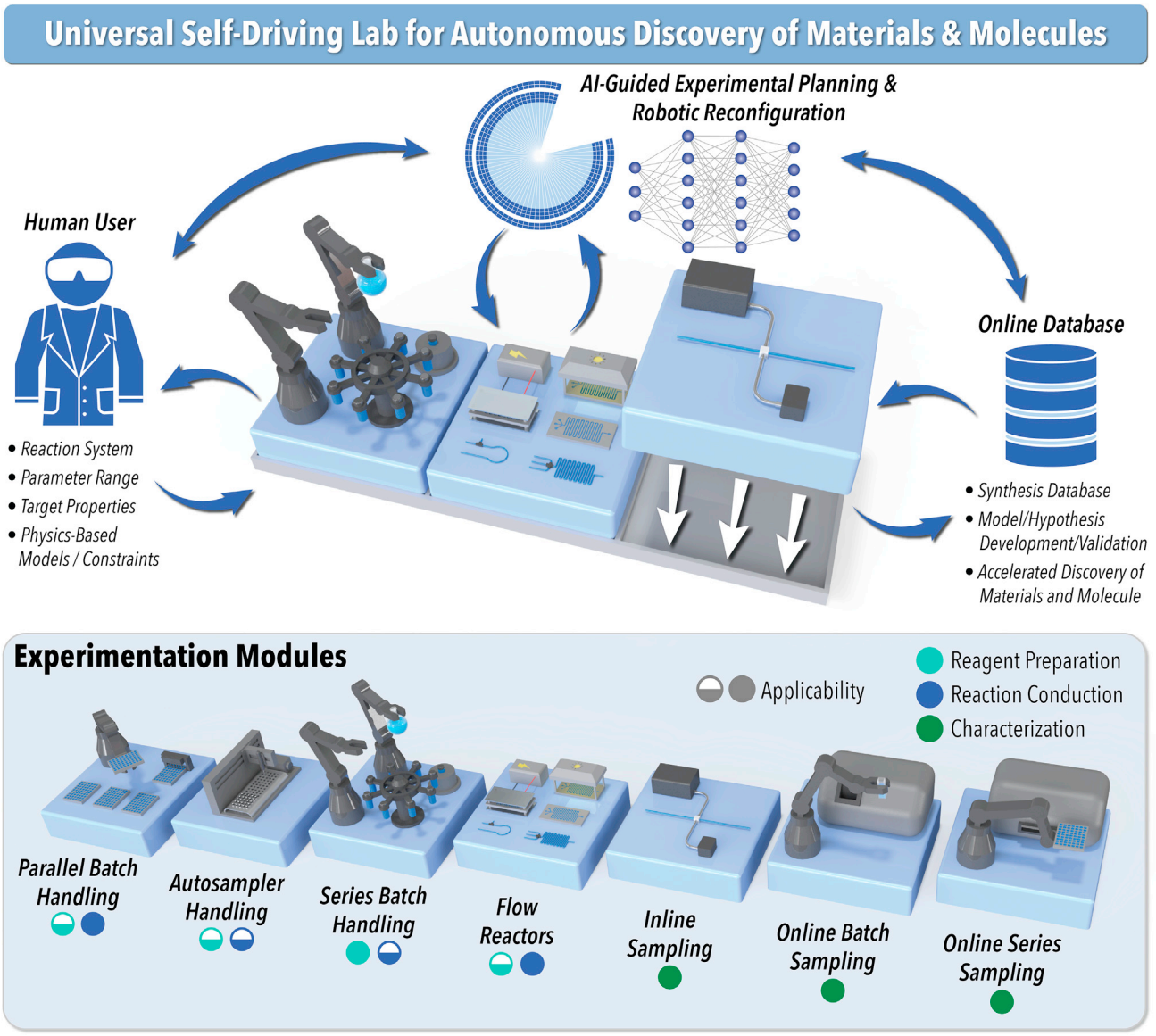

Figure 1. Schematic illustration of a modular autonomous experimentation strategy for a universal self-driving lab

Automated experimentation modules can be assembled as needed, and data generation is both fed into decision-making algorithms and archived into a larger communal database. The human user defines the constraints from which the system operates and its ultimate target properties or objective.

illustrated in Figure 1, would improve the rate of hybrid system development and lower the barrier for entry among researchers. Furthermore, this modular approach to platform development would provide a direct precursor to autonomous device fabrication and optimization.[10] Sharing of datasets generated with these unified platforms would then provide large quantities of transferable experimental information, leading to higher performing AI models and algorithms and greater mastery of the chemical world.

Unifying the direction of research and shifting the current standards of the field will require a focused effort and cooperation across academia and industry. One promising step in this direction is the recently established *Acceleration Consortium* hosted at the University of Toronto. The output of this association and similar future

establishments will help reveal the efficiencies of self-driven experimentation platforms to a larger audience. While industry will always have strong incentives to maintain proprietary information about novel materials and molecules, they can benefit from collaborations with academic researchers to develop an accessible modular experimentation core from which the self-driving platforms may be built, applied, and expanded upon.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the National Science Foundation (Award # 1940959) and the UNC Research Opportunities Initiative (UNC-ROI) grant.

DECLARATION OF INTERESTS

The lead contact is a Global Member of the Acceleration Consortium.

REFERENCES

1. Burger, B., Maffettone, P.M., Gusev, V.V., Aitchison, C.M., Bai, Y., Wang, X., Li, X., Alston, B.M., Li, B., Clowes, R., et al. (2020). A mobile robotic chemist. Nature *583*, 237–241.

2. Chan, E.M., Xu, C., Mao, A.W., Han, G., Owen, J.S., Cohen, B.E., and Milliron, D.J. (2010). Reproducible, high-throughput synthesis of colloidal nanocrystals for optimization in multidimensional parameter space. Nano Lett. *10*, 1874–1885.

3. Steiner, S., Wolf, J., Glatzel, S., Andreou, A., Granda, J.M., Keenan, G., Hinkley, T., Aragon-Camarasa, G., Kitson, P.J., Angelone, D., et al. (2019). Organic synthesis in a modular robotic system driven by a chemical programming language. Science *363*, eaav2211.

4. Bédard, A.C., Adamo, A., Aroh, K.C., Russell, M.G., Bedermann, A.A., Torosian, J., Yue, B., Jensen, K.F., and Jamison, T.F. (2018). Reconfigurable system for automated optimization of diverse chemical reactions. Science *361*, 1220–1225.

5. Coley, C.W., Thomas, D.A., Lummiss, J.A.M., Jaworski, J.N., Breen, C.P., Schultz, V., Hart, T., Fishman, J.S., Rogers, L., Gao, H., et al. (2019). A robotic platform for flow synthesis of organic compounds informed by AI planning. Science *365*, eaax1566.

6. Epps, R.W., Bowen, M.S., Volk, A.A., Abdel-Latif, K., Han, S., Reyes, K.G., Amassian, A., and Abolhasani, M. (2020). Artificial Chemist: An Autonomous Quantum Dot Synthesis Bot. Adv. Mater. *32*, e2001626.

7. Buitrago Santanilla, A., Regalado, E.L., Pereira, T., Shevlin, M., Bateman, K., Campeau, L.-C., Schneeweis, J., Berritt, S., Shi, Z.-C., Nantermet, P., et al. (2015). Organic chemistry. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. Science *347*, 49–53.

8. Volk, A.A., Epps, R.W., and Abolhasani, M. (2021). Accelerated Development of Colloidal Nanomaterials Enabled by Modular Microfluidic Reactors: Toward Autonomous Robotic Experimentation. Adv. Mater. *33*, e2004495.

9. Salley, D., Keenan, G., Grizou, J., Sharma, A., Martín, S., and Cronin, L. (2020). A nanomaterials discovery robot for the Darwinian evolution of shape programmable gold nanoparticles. Nat. Commun. *11*, 2771.

10. MacLeod, B.P., Parlane, F.G.L., Morrissey, T.D., Häse, F., Roch, L.M., Dettelbach, K.E., Moreira, R., Yunker, L.P.E., Rooney, M.B., Deeth, J.R., et al. (2020). Self-driving laboratory for accelerated discovery of thin-film materials. Sci. Adv. *6*, eaaz8867.